

Université de Montréal

Influence des expansions de territoire sur la capacité des approches en génomique du paysage d'identifier les gènes adaptatifs

Par Paul Mayrand

Département de sciences biologiques

Faculté des arts et sciences

Mémoire présenté

en vue de l'obtention du grade de M.Sc.

en sciences biologiques

Janvier 2017

© Paul Mayrand, 2017

Résumé:

Avec les changements climatiques et les perturbations humaines, de nombreuses espèces changent leurs aires de répartition. Dans ce contexte, l'identification de loci potentiellement adaptatifs chez ces populations en expansion est importante pour mieux comprendre le potentiel évolutif et la capacité d'envahissement de ces espèces. Toutefois, chez les espèces en expansion de territoire, tel le dendroctone du pin, les processus démographiques comme le surf d'allèles peuvent résulter en des patrons spatiaux de variation génétique neutre qui imitent ceux issus des processus adaptatifs. Ce phénomène gonfle le taux de faux-positif d'identification des loci adaptatifs et confond ainsi les méthodes de génomique du paysage. Dans le cadre de ce projet de maîtrise, j'ai étudié le développement des structures génétiques neutres et adaptatives lors d'une expansion de territoire. Je me suis attardé particulièrement sur l'influence de différentes conditions démographiques sur le taux de loci neutres qui imitent les patrons spatiaux de ceux adaptatifs, en menant une revue de littérature et en utilisant le système épidémique du dendroctone du pin pour paramétrer un modèle de simulations. J'ai simulé les dynamiques démographiques et génétiques des populations de dendroctones à l'aide du modèle de simulation génétique explicitement spatial CDmetaPOP. J'ai analysé les conséquences de trois facteurs sur le taux de faux-positif : 1) la capacité de dispersion; 2) le moment d'échantillonnage durant l'expansion de territoire; et 3) la force de sélection agissant sur le locus adaptatif de référence. J'ai démontré qu'une combinaison de faible capacité de dispersion, faible sélection et un échantillonnage tôt au début de l'expansion contribuent à un plus grand taux de faux-positif, alors qu'une forte capacité de dispersion conduit à de plus faibles taux de faux-positif. Lorsque les méthodes de génomique du paysage sont utilisées dans ces conditions, elles risquent d'avoir un haut taux de loci neutres identifiés comme adaptatifs et doivent donc être interprétées avec prudence. La situation démographique complexe que présente le système actuel du dendroctone (et d'autres espèces irruptives et envahissantes) rend l'identification d'allèles adaptatifs plus difficile. Les résultats de ce projet encouragent l'incorporation de ces processus démographiques dans les méthodes de génomique du paysage.

Mots clés:

dendroctone du pin, expansion du territoire, génomique du paysage, CDmetaPOP, variation génétique adaptative, variation génétique neutre, dispersion, sélection

Abstract:

Under the actual climate changes and human perturbations global context, many species are altering their geographic range. The identification of putatively adaptive loci in those expanding populations is thus important to better understand evolutionary potential and invasiveness of these species. However, in irruptive species undergoing rapid expansion, such as the mountain pine beetle (MPB), demographic processes such as allele surfing can result in spatial patterns of neutral genetic variation that can mimic those that result from adaptive processes. This phenomenon inflates the false discovery rate of adaptive loci and thus confounds landscape genomics methods. In this thesis, I studied the development of neutral and adaptive genetic structure during a range expansion. I investigated precisely how different demographic conditions influence the rate of neutral loci mimicking the spatial patterns of adaptive loci, doing a literature review and using the mountain pine beetle outbreak system to parametrize a simulation model. I simulated the demographic and population genetic dynamics of mountain pine beetle populations undergoing range expansion using the spatially explicit, individual-based genetic model CDmetaPOP. I examined the consequences of three factors on the false discovery rate: 1) species dispersal capacity; 2) timing of sampling during the course of the expansion; and 3) the strength of selection on adaptive reference loci. I found that a combination of weak dispersal capacity, weak selection, and early sampling during expansion results in the highest rate of false positive, while strong dispersal was responsible for lower rates of false positive. Used under these conditions of dispersal capacity, strength of selection and sampling timing, landscape genomics models risk elevated false discovery rates of adaptive loci and must be interpreted cautiously. Complex demography in the current MPB system (and other irruptive and invasive species) makes identification of adaptive loci challenging. Results from this project clearly demonstrate that there is a need for further method development to include these directional demographic processes in the field of landscape genomics.

Keywords:

mountain pine beetle, range expansion, landscape genomics, CDmetaPOP, adaptive genetic variation, neutral genetic variation, dispersal, selection

TABLE DES MATIÈRES

Chapitre 1 : Introduction du mémoire.....	1
1. Introduction.....	1
2. Questions et objectifs de recherche.....	2
3. Pertinence.....	3
4. Cadre expérimental	3
4.1. Approche par modélisation	3
4.2. CDMetaPOP	4
5. Revue de la littérature.....	5
5.1. L'expansion de territoire	5
5.1.1. Quelques exemples d'expansion de territoire	5
5.1.2. Distinction entre expansion démographique et expansion de territoire.....	6
5.1.3. Conséquences génétiques d'une expansion de territoire.....	7
5.2. Cadre théorique.....	12
5.2.1. Génétique du paysage	12
5.2.2. Génomique du paysage	12
5.2.3. Limites et problématique	13
5.3. Le dendroctone du pin	15
5.3.1. Résumé de la biologie.....	16
5.3.2. Capacité de dispersion	17
5.3.3. Épidémies	17
5.3.4. Conséquences de l'épidémie en cours	18
5.3.5. Structure génétique du dendroctone du pin.....	19
6. Hypothèses.....	19
Chapitre 2: How range expansion can confound landscape genomic approaches to identify loci under selection	20
Abstract	21
Introduction.....	22
Mountain pine beetle range expansion	23
Methods	25
Modelling framework.....	25
Dispersal	25

Reproduction	28
Selection	28
Simulation scenarios.....	28
Identifying outlier loci.....	30
Model validation.....	31
Results	32
Simulations	32
Model validation.....	32
Identifying outlier loci.....	34
Fixation of the adaptive locus	38
Discussion	39
Dispersal	40
Timing of sampling	40
Selection	41
Models of selection	41
Conclusions.....	42
Acknowledgements	43
Supporting information	44
Chapitre 3 : Conclusion.....	48
Bibliographie	52

LISTE DES TABLEAUX

Table I : Summary of the scenarios combinations of parameters	30
Table II : Summary of the ANOVA results	38
Table S.I : Simulation parameters summary	44

LISTE DES FIGURES

Figure 1 : Dispersal model	27
Figure 2 : PCA plots based on three distinct simulations varying in dispersal capacity and at two different time steps	33
Figure 3 : Example of a scatter plot illustrating the relationship between allele frequency and geographic position	34
Figure 4 : Frequency distribution of absolute neutral allele slopes for three different dispersal capacities at generation 150	35
Figure 5 : Boxplots of neutral alleles with a higher slope value than the mean adaptive slope value, through time and for the three dispersal capacities and three strengths of selection ..	36
Figure 6 : Mean adaptive slope over 50 replicates through time for three different dispersal capacities and at a weak strength of selection	39

LISTE DES SIGLES

MPB : Mountain pine beetle (*Dendroctone du pin*)

REMERCIEMENTS

J'aimerais remercier un certain nombre de personnes qui ont contribué à l'accomplissement de mon projet de maîtrise et de ce mémoire, et pour qui j'en serai toujours très reconnaissant.

Tout d'abord, j'aimerais remercier mille fois mon superviseur, Patrick James, qui a appuyé et guidé mes démarches depuis le début. Ses encouragements, ses bons conseils et son esprit critique ont contribué à mon développement professionnel et scientifique. Je lui serai toujours reconnaissant de m'avoir encouragé à participer à une grande diversité d'événements scientifiques, où j'ai eu l'occasion d'acquérir de l'assurance et de la confiance en moi lors des présentations et échanges avec des collègues. Patrick a toujours su comment mettre mes travaux en valeur et me faire découvrir de nouvelles facettes de la recherche scientifique et du potentiel de mes efforts : il est un véritable modèle pour quiconque souhaite poursuivre une carrière en recherche.

Autant de remerciements doivent être faits à Élise Filotas, ma co-directrice, qui était toujours disponible pour me rencontrer et discuter de mon projet. Élise a apporté une perspective différente et importante, notamment au niveau de la modélisation et de l'utilisation de mon simulateur. Tout au long du processus de recherche, elle s'est assurée de la qualité de mes travaux grâce à des commentaires éclairés et constructifs qui m'ont constamment invité à la réflexion et au perfectionnement de mes aptitudes. Ainsi, pour toute son aide précieuse, je remercie chaleureusement Élise.

Les autres membres de mon comité conseil de maîtrise et du jury de mon mémoire, Jesse Shapiro, Simon Joly et Virginie Millien, ont également contribué au bon développement de mon mémoire et de mon projet, grâce à leur soutien et leurs commentaires pertinents. Jesse Shapiro m'a aidé à développer davantage la perspective génétique de mon projet, en m'indiquant de bonnes pistes de recherche ainsi qu'en dévoilant les faiblesses de mon projet à cet égard. Les commentaires de Simon Joly, quant à eux, m'ont permis de revoir plusieurs volets de mes recherches, notamment en ce qui concerne mon modèle de sélection. Simon a également contribué à l'élaboration de ma réflexion sur la portée de mes résultats et ce qu'ils représentent vraiment. Enfin, Virginie Millien m'a spécialement aidé à mettre mon projet dans une optique plus globale, afin de le rendre plus accessible aux lecteurs de tous horizons et afin d'en prouver la pertinence.

J'aimerais aussi remercier tous mes collègues de laboratoire, que je côtoie depuis près de 3 ans pour la plupart, et qui ont apporté un regard avisé sur tous les volets de mes recherches. Merci à Julian Wittische pour avoir souvent remis des portions entières de mes recherches en question et pour avoir introduit de nouveaux concepts et réflexions cruciaux à mon projet. J'ai très apprécié sa franche camaraderie et nos péripéties à Edmonton. Pour sa bonne humeur et son support moral, je veux également remercier Olivier Pontbriand-Paré, qui a traversé les mêmes épreuves en effectuant son projet de maîtrise en même temps que moi. J'ai eu beaucoup de plaisir en sa compagnie et il a su contribué au développement de mon projet grâce à son esprit critique aiguisé. Simon Legault mérite tout autant mes remerciements pour ses conseils judicieux et sa grande expérience en enseignement et en communication scientifique. Ce fut un réel plaisir d'assister à plusieurs conférences scientifiques en sa compagnie. Merci aussi à Louis-

Étienne Robert pour son inépuisable savoir en statistiques et en programmation, auquel j'ai eu maintes fois recours. Louis-Étienne a toujours été très généreux, aidant quiconque au laboratoire rencontrant des difficultés. Un grand merci aussi à Julie Marleau, pour la joie qu'elle apporte au laboratoire et pour son aide dans les tâches administratives. Enfin, j'aimerais remercier Soufiane Tahir pour son aide précieuse au tout début du projet et qui travailla fort à mes côtés pour comprendre le fonctionnement du simulateur. J'aimerais remercier d'autres collègues qui ont eu un passage plus bref au laboratoire ou que je n'ai pas pu côtoyer suffisamment longtemps, mais qui ont eu l'occasion de poser sur mon projet un regard nouveau. Ainsi, merci à Patricia Sanae Suji, Ronan Marrec, Camille Marier-Desroches, Mathieu Neau, Chloé Frédette, Colin Garroway, Élise Saint-Pierre, Arnald Marcer, Anaïs Boa, Jennifer Sauri, Jeremy Larroque et Roger-Michael Deschênes.

Je tiens aussi à remercier d'autres contributeurs majeurs à mon projet de maîtrise, qui m'ont aidé sur un aspect plus technique. Ainsi, je serai éternellement reconnaissant envers Erin Landguth de la University of Montana pour avoir répondu à toutes mes questions de néophyte concernant le fonctionnement de ses simulateurs, CDPop et CDmetaPOP. Erin a toujours été très généreuse de son temps pour s'assurer que je comprenne bien les fonctionnalités de son programme et s'est toujours montrée intéressée par mes travaux, me conseillant même sur certaines façons de procéder dans mes recherches. Également, je souhaite remercier toute la merveilleuse équipe derrière Calcul Québec, une ressource de calcul informatique de pointe qui s'est prouvée indispensable pour mes simulations et analyses. Notamment, son personnel dévoué et plus que qualifié m'ont plus d'une fois sauvé la mise lorsque j'utilisais leurs ressources. Ainsi, je suis très reconnaissant envers Julie Faure-Lacroix et Félix-Antoine Fortin pour leur support indispensable.

Je tiens aussi à remercier le projet TRIA-Net pour l'aide financière de ces deux dernières années et pour m'avoir permis de me rendre à deux reprises à une assemblée générale à Edmonton afin d'y rencontrer des collègues et autres chercheurs. Merci au programme de formation en Modélisation de la complexité de la forêt (MCF) qui m'a lui aussi apporté un soutien financier, en m'offrant également des formations pertinentes en modélisation écologique. Finalement, j'aimerais donner mes remerciements au Centre de la science de la biodiversité du Québec (CSBQ) et au Centre d'étude de la forêt (CEF) pour leurs bourses de voyage qui me permirent d'assister à un congrès international en Caroline du Nord.

Finalement, je souhaite remercier tous ceux qui sont présents dans ma vie et qui m'ont aidé à traverser cette épreuve, m'appuyant durant les moments plus difficiles. Ma famille et mes amis ont été très précieux pour moi et je suis certain qu'en bout de ligne ils ont eu des répercussions positives sur mon projet de maîtrise. Je serai toujours reconnaissant envers mes parents, Nicole et Michel, pour leur soutien et leurs encouragements indéfectibles. Ils m'ont appris beaucoup de choses de la vie et c'est grâce à eux qu'aujourd'hui je me suis rendu jusque-là. Je tiens à remercier tous mes grands *chums* qui ont égaillé mes soirées et mes fins de semaine : merci à Philippe, Simon L., Gabriel, Jean-Christophe, Jules, Simon G., Ludwig et Valentin. Enfin, un très grand merci à ma copine Gabrielle, que j'aime plus que tout et qui est toujours présente pour moi.

Chapitre 1 : Introduction du mémoire

1. Introduction

La fragmentation des habitats (Wilson et al., 2016) et les changements climatiques (Thornton, Ericksen, Herrero, & Challinor, 2014) ont de multiples impacts sur les fonctions des écosystèmes à l'échelle de la planète (Haddad et al., 2015). En réponse à ces changements importants, les espèces recourent à trois stratégies: l'acclimatation (plasticité phénotypique), l'adaptation via la sélection, et la colonisation de nouveaux habitats (Hansen, Olivieri, Waller, & Nielsen, 2012). Cette dernière stratégie peut avoir des répercussions importantes puisque les populations en expansion bouleversent la structure des communautés écologiques et peuvent altérer les fonctions écosystémiques qu'elles soutiennent (Chuang & Peterson, 2016).

L'adaptation aux conditions locales peut faciliter les expansions de territoire en permettant aux populations de s'adapter à des pressions de sélection différentes d'un nouvel habitat (Tiffin & Ross-Ibarra, 2014). Dans le contexte de changements globaux que nous connaissons actuellement, l'étude de la variation génétique sous sélection est de plus en plus populaire pour repérer rapidement les gènes qui peuvent contribuer au succès d'expansion et de colonisation des espèces envahissantes et opportunistes (Landguth & Balkenhol, 2012; Moran & Alexander, 2014). Étudier la structure génétique de telles populations est un moyen efficace pour comprendre ce phénomène d'expansion de territoire et évaluer le potentiel d'expansion d'une espèce.

L'expansion de territoire est un phénomène démographique complexe qui a de grands impacts sur la structure génétique d'une population. En effet, plusieurs processus génétiques sont à l'œuvre comme le flux génique, la dérive génétique et la sélection naturelle. Le flux génique consiste en un échange de gènes entre des populations ou sous-populations distinctes, via la migration et la reproduction d'un ou plusieurs individus d'un groupe à l'autre (Hartl & Clark, 1997). La dérive génétique concerne la fluctuation d'une génération à l'autre des fréquences alléliques d'une population, qui est le résultat de la sélection aléatoire des gamètes parmi l'ensemble de ceux disponibles lors de la formation d'un embryon (Ford, 1975). Enfin, la sélection naturelle correspond aux conditions environnementales qui favorisent le maintien ou la suppression de caractères liés à la survie ou la reproduction d'un individu (Hartl & Clark,

1997). Ces trois mécanismes influencent à la fois la variation génétique neutre, qui concerne la diversité des gènes qui n'affectent pas la capacité de survie ou de reproduction d'un individu, et celle adaptative, qui concerne la diversité des gènes qui influencent la capacité de survie ou de reproduction d'un individu (Excoffier, Foll, & Petit, 2009). Étudier la variation génétique sous sélection est un objectif essentiel pour mieux comprendre les mécanismes derrière les réponses évolutives d'une expansion de territoire (Hansen et al., 2012). La génomique du paysage est une branche de l'écologie relativement nouvelle et intégrative qui tente d'identifier des loci sous sélection en analysant la corrélation entre les fréquences alléliques et des variables environnementales (Schwartz, Luikart, McKelvey, & Cushman, 2010). Cependant, plusieurs méthodes statistiques actuellement employées pour identifier des gènes potentiellement adaptatifs ne prennent pas en compte toute la complexité démographique propre aux expansions de territoire. En particulier, une forte structure génétique de la population et une série d'effets fondateurs au front de l'expansion peuvent conduire les loci neutres à imiter les patrons spatiaux issus des loci adaptatifs (Excoffier & Ray, 2008; Vasemägi, 2006). Conséquemment, ces méthodes risquent de relever des résultats fallacieux lorsqu'elles sont utilisées dans des situations démographiques particulières comme une expansion de territoire, en identifiant des gènes neutres comme étant potentiellement adaptatifs (François, Martins, Caye, & Schoville, 2016; Parisod & Holderegger, 2012). Les taux d'erreur peuvent ainsi être parfois très importants (Forester, Jones, Joost, Landguth, & Lasky, 2016; Frichot, Schoville, de Villemereuil, Gaggiotti, & François, 2015) et il devient donc nécessaire de bien connaître les facteurs démographiques, temporels et spatiaux qui biaisent ces méthodes d'association environnementale.

2. Questions et objectifs de recherche

L'objectif de ce projet de maîtrise est d'améliorer nos connaissances sur les processus qui créent les structures génétiques neutres et adaptatives spatiales chez les espèces en expansion du territoire. En particulier, ce projet s'intéresse au cas du dendroctone du pin (*Dendroctonus ponderosae*) et vise à relever la façon par laquelle les interactions entre la capacité de dispersion, la force de sélection et le moment d'échantillonnage influencent le développement ou la détection de la structure génétique neutre et adaptative, en utilisant un modèle de simulation spatiale. Ce modèle servira également à examiner les conséquences de

l'expansion territoriale sur notre capacité à identifier les allèles sous sélection et à les distinguer des allèles neutres.

3. Pertinence

Étudier la structure génétique d'une population en expansion représente un intérêt particulier pour la recherche théorique en écologie et en évolution, car ces expansions constituent des expériences écologiques et évolutives naturelles survenues très récemment (Bock et al., 2015). Aussi, cela permet d'approfondir les connaissances sur les facteurs déterminants associés aux phénomènes d'épidémies et d'espèces ravageuses (Goodsman, Cooke, Coltman, & Lewis, 2014). Étudier la structure génétique adaptative permet ainsi de cibler les gènes qui conduisent au succès de colonisation et d'expansion et renseigne également sur la future progression d'une telle expansion (Moran & Alexander, 2014).

Également, dans l'étude des invasions biologiques, tenir compte des particularités génétiques dynamiques de l'espèce envahissante est primordial pour améliorer les modèles prédictifs (Barrett, 2015; Cristescu, 2015; Sakai et al., 2001). En effet, le développement de la structure des fréquences alléliques adaptatives a souvent un effet déterminant sur la colonisation (Lehe, Hallatschek, & Peliti, 2012). Enfin, relever la structure génétique d'une population renseigne également sur son histoire démographique récente et ancienne : il est ainsi possible d'identifier la présence d'une expansion de territoire (Khatchikian et al., 2015), de retracer les routes de migrations empruntées (Taberlet, Fumagalli, Wust-Saucy, & Cosson, 1998) ou même encore de décrire la situation démographique et génétique d'une population source éteinte depuis des centaines, voire des milliers d'années (Liu, Prugnolle, Manica, & Balloux, 2006).

4. Cadre expérimental

4.1. Approche par modélisation

Afin d'étudier le développement de structures génétiques spatiales sur un territoire lors d'une expansion, ce projet utilise l'approche de modélisation par simulations. Examiner des phénomènes écologiques via la modélisation permet d'explorer la dynamique des systèmes sur de vastes échelles de temps et d'espace (Epperson et al., 2010). La modélisation est particulièrement utile pour reproduire des processus qui varient selon la configuration des paysages et les conditions locales précises, qui sont difficilement reproductibles dans la nature

(Larroque, Ruet, Vandell, & Devillard, 2016). Cela permet aussi un contrôle absolu sur les paramètres à l'étude, en les faisant varier selon les besoins (Epperson et al., 2010). Utiliser un modèle de simulation permet également d'examiner le fonctionnement d'un système, en le comparant à des situations connues d'expansion de territoire et de structuration génétique et de comprendre des éléments clés de la dynamique des populations (Hoban, Bertorelle, & Gaggiotti, 2012). De plus, dans le cadre de mon projet, la modélisation de loci adaptatifs, identifiables à priori dans mes données génétiques, permet d'analyser la performance des méthodes d'identification de gènes potentiellement adaptatifs.

Mon étude utilise un modèle de simulation stochastique servant à inclure la variabilité propre aux événements démographiques et évolutifs (Bridle, Polechová, Kawata, & Butlin, 2010). La variation stochastique incluse dans les processus modélisés permet de représenter certains phénomènes plus rares (ex : le surf d'allèles). De plus, la modélisation stochastique permet de reproduire des phénomènes complexes pouvant difficilement être modélisés par des fonctions mathématiques connues (Epperson et al., 2010). La modélisation stochastique exige la production d'un grand nombre de simulations indépendantes. L'analyse statistique de l'ensemble des simulations produites permet d'identifier les patrons dominants et leur variabilité. L'analyse doit porter sur un échantillon suffisamment grand de simulations, sans quoi une trop forte importance statistique pourraient être données à certains phénomènes stochastiques rares (Goodsman et al., 2014). Il existe présentement plusieurs modèles de simulations stochastiques pouvant représenter simultanément les processus génétiques, démographiques et environnementaux (Hoban et al., 2012). Ces modèles peuvent ainsi répondre à une panoplie de questions en écologie moléculaire (Hoban, 2014) et représentent donc un cadre riche et efficace pour conduire des simulations d'expansion de territoire.

4.2. CDMetaPOP

CDMetaPOP fait justement partie de ces simulateurs génétiques polyvalents qui modélisent des processus démographiques et génétiques complexes dans un contexte spatial (Landguth, Bearlin, Day, & Dunham, 2016). Né tout récemment de la fusion entre les simulateurs CDPOP (Landguth & Cushman, 2010) et CDFISH (Landguth, Muhlfeld, & Luikart, 2012), CDMetaPOP permet de suivre la progression spatio-temporelle de la structure génétique d'une population à l'échelle des individus. En effet, avec ce simulateur, il est possible de

surveiller de génération en génération les échanges génétiques entre les individus et la dispersion des descendants à travers un territoire conçu sur mesure. L'ancêtre de ce simulateur (c.-à-d. CDPOP) fut utilisé à quelques reprises pour des analyses de sensibilité aux perturbations climatiques et anthropiques (Castillo, Epps, Davis, & Cushman, 2014; Dileo, Rouse, Dávila, & Loughheed, 2013; Forester et al., 2016; Wasserman, Cushman, Shirk, Landguth, & Littell, 2012), mais également pour aborder des questions théoriques de génétique des populations en simulant des populations animales fictives (Cushman, Shirk, & Landguth, 2012; Cushman, Shirk, & Landguth, 2013; Landguth & Balkenhol, 2012; Erin L. Landguth, Bradley C. Fedy, et al., 2012). CDMetaPOP est maintenant un produit bien plus sophistiqué que la première version de CDPOP, qui avait à l'époque été évaluée négativement par la classification des simulateurs de Hoban et al. (2012) : il est dorénavant possible de simuler des scénarios démographiques complexes, d'inclure la sélection naturelle (Landguth, Cushman, & Johnson, 2012) et de finement paramétrer la capacité de dispersion. Toutes ces caractéristiques font de CDMetaPOP un simulateur de choix pour modéliser une situation démographique complexe telle une expansion de territoire.

Ainsi, c'est avec cette approche par modélisation que je pourrai simuler des expansions de territoire et tester l'influence de la capacité de dispersion, de la force de sélection et du temps sur notre aptitude à identifier correctement les loci sous sélection. Plus précisément, je me pencherai sur l'influence de ces facteurs sur le taux de faux-positif et tenterai de mieux comprendre la structuration génétique neutre durant une expansion de territoire.

5. Revue de la littérature

Cette section dresse un portrait de la littérature permettant de définir les concepts clés et le cadre théorique dans lesquels s'insère mon projet de recherche. En particulier, je me penche sur trois champs de connaissances : l'expansion du territoire (problématique), la génétique et la génomique du paysage (cadre théorique), et le dendroctone du pin (cas d'étude).

5.1. L'expansion de territoire

5.1.1. *Quelques exemples d'expansion de territoire*

Depuis un peu plus d'une décennie, plusieurs cas d'expansion de territoire favorisés par les changements climatiques ont été relevés chez plusieurs populations naturelles appartenant à

divers taxons. C'est le cas par exemple d'une espèce de crabes violonistes (*Uca pugnax*; Sanford, Holzman, Haney, Rand, & Bertness, 2006), de *Centrostephanus rodgersii*, un oursin australien (Ling, Johnson, Frusher, & King, 2008) ou encore de plusieurs espèces de demoiselles (Hassall, Thompson, & Harvey, 2009; Therry, Lefevre, Bonte, & Stoks, 2014) et de papillons (Braschler & Hill, 2007; J.K. Hill, Thomas, & Blakeley, 1999). Ces expansions reflètent d'ailleurs la situation rencontrée avec la période post-glaciale survenue il y a quelques milliers d'années (Jane K. Hill, Griffiths, & Thomas, 2011).

Bien entendu, l'entière expansion de territoire n'est pas forcément déclenchée par les changements climatiques : les introductions d'espèces sont responsables pour bon nombre d'entre eux, comme c'est le cas avec le gobie à taches noires (*Neogobius melanostomus*; Raby, Gutowsky, & Fox, 2010), le crapaud buffle (*Rhinella marina*; Brown, Phillips, & Shine, 2015), le moineau domestique (*Passer domesticus*; Liebl & Martin, 2012), la coccinelle asiatique (*Harmonia axyridis*; Lombaert et al., 2014) et le chèvrefeuille du Japon (*Lonicera japonica*; Kilkenny & Galloway, 2013). Enfin, certaines expansions de territoire surviennent dans le cadre naturel des processus évolutifs et écologiques (telle la recolonisation d'un territoire auparavant occupé), comme il peut être actuellement observé avec le castor (*Castor fiber*; Hartman, 1997), la mélitée du plantain (*Melitaea cinxia*; Haag, Saastamoinen, Marden, & Hanski, 2005) et le pin tordu (*Pinus contorta*; Cwynar & MacDonald, 1987).

5.1.2. Distinction entre expansion démographique et expansion de territoire

Une expansion de territoire est toujours caractérisée par une augmentation du nombre d'individus et est ainsi souvent associée à une simple expansion démographique (Excoffier et al., 2009). Il faut se pencher sur les conséquences de l'expansion sur la structure génétique de la population pour bien saisir les distinctions entre ces deux formes d'expansion.

En règle générale, avec un plus grand nombre d'individus dans une population, les chances de créer de la variabilité génétique augmente avec le taux d'apparition de mutations (Ford, 1975; Hague & Routman, 2016). Cette croissance, expliquée par exemple par une baisse de la pression de sélection due à une perturbation de l'habitat, fait donc en sorte que la population se munit d'une variabilité génétique favorable à l'adaptation face à de nouvelles conditions environnementales (Willi, Van Buskirk, & Hoffman, 2006). Ces conditions favorables à l'évolution peuvent permettre aux individus de coloniser des milieux à l'extérieur de leur

présente aire de répartition, où les conditions sont plus ou moins différentes. Ainsi, dans cette situation, une expansion de territoire succède une expansion démographique.

Comparée à une situation stable, une population qui connaît une croissance démographique devient moins sensible aux fluctuations des allèles engendrées par la dérive génétique, ce qui stabilise les fréquences alléliques (Kimura & Crow, 1963). Un excès d'allèles rares est également caractéristique d'une augmentation d'individus (Maruyama & Fuerst, 1984), ainsi qu'un excès d'homozygotes (Watterson, 1986). Une fois le processus d'expansion territoriale enclenché, le contexte spatial occasionne des mécanismes génétiques beaucoup plus complexes que durant une simple expansion démographique, qui peuvent avoir par exemple d'importants rôles dans le succès d'invasion d'une espèce (Drury, Drake, Lodge, & Dwyer, 2007; Garnier, Roques, & Hamel, 2012). Examinons maintenant en détails les conséquences sur la structure génétique d'une expansion de territoire.

5.1.3. Conséquences génétiques d'une expansion de territoire

Je vais maintenant décrire quatre facteurs qui influencent la structure génétique d'une population lors d'une expansion de territoire : a) la distribution spatiale de la population, b) le mécanisme de surf d'allèles, c) la capacité de dispersion et d) le temps écoulé depuis le début de l'expansion.

5.1.3.a. Cœur versus périphérie

Une population n'est pas démographiquement distribuée de manière homogène sur le territoire qu'elle occupe. En effet, les individus sont généralement concentrés dans la zone centrale de leur aire de répartition et sont plus diffus en périphérie (Phillips, Brown, & Shine, 2010) : ceci amène une dynamique populationnelle particulière qui a des répercussions sur la structure génétique (Chuang & Peterson, 2016). Étant ainsi isolés de la vaste majorité de leurs congénères, les individus en périphérie subissent différemment l'effet de la dérive génétique et tendent donc à se différencier génétiquement (Hallatschek, Hersen, Ramanathan, & Nelson, 2007). D'ailleurs, ils font également face à des conditions environnementales différentes de celles rencontrées au cœur de la population, impliquant une différenciation génétique supplémentaire dirigée par la sélection naturelle (Zenni & Hoban, 2015). De plus, étant donné la grande différence de compétition intra-spécifique entre le cœur et la périphérie, des traits différents sont sélectionnés : au cœur, ce sont les traits qui favorisent la compétitivité (Phillips

et al., 2010) et en périphérie, ceux qui accroissent la reproduction et l'utilisation de ressources (Reznick, Bryant, & Bashey, 2002). Tout de même, le centre et la périphérie restent liés, mais selon un système de source-puit, où le flux génique s'effectue principalement de manière unidirectionnelle vers la périphérie (Kirkpatrick & Barton, 1997). Si ce flux est suffisamment important, les groupes d'individus en périphérie peuvent voir leurs nouvelles adaptations et différenciations génétiques être diluées et même éliminées du bassin génétique (Kirkpatrick & Barton, 1997).

Cette structure d'une population stable est également retrouvée dans une population en expansion du territoire et se développe de manière dynamique avec le temps et la progression de l'expansion (Phillips et al., 2010). Durant une expansion de territoire, on nomme « front de l'expansion » la zone où la plus récente colonisation d'un nouvel habitat a eu lieu (Chuang & Peterson, 2016). Ce front correspond à la zone périphérique où la densité en individus est beaucoup plus faible et où la dérive génétique peut avoir de grands impacts (Excoffier et al., 2009). De plus, étant donné la nature dynamique d'une expansion, ce front progresse dans l'espace, laissant derrière lui des sous-populations d'individus qui connaissent des expansions démographiques et qui s'intègrent alors au cœur de la population (Chuang & Peterson, 2016). Ce front de l'expansion en plein mouvement provoque un continuel rééchantillonnage des nouveaux colonisateurs : il y a ainsi une série d'effets fondateurs chaque fois que l'expansion progresse dans l'espace (Slatkin & Excoffier, 2012). Ceci a pour principale conséquence une perte graduelle de la diversité génétique, qui procède au fil des nouveaux goulets d'étranglement génétique (Hallatschek & Nelson, 2008). Bref, les individus sur le front d'une expansion expérimentent les processus génétiques à un niveau très différent de ceux qui se trouvent au centre du territoire (Hallatschek & Nelson, 2009).

5.1.3.b. Surf d'allèles

Récemment, l'étude des expansions de territoire a mis en évidence un phénomène particulier causé par le constant rééchantillonnage du matériel génétique sur le front de l'expansion : le surf d'allèles (« *allele surfing* ») (Klopfstein, Currat, & Excoffier, 2006). En effet, les études théoriques à ce sujet ont proliféré ces dernières années (Burton & Travis, 2008a; Excoffier et al., 2009; Hallatschek et al., 2007; Hallatschek & Nelson, 2008). Le surf d'allèles correspond à l'atteinte de hautes fréquences alléliques au front de l'expansion, que ces allèles soient neutres, avantageux ou même délétères (Travis et al., 2007). Les nouvelles mutations

ainsi que la variation génétique existante sont toutes deux sujettes au surf d'allèles (Excoffier et al., 2009), quoique la présence et l'importance de ce phénomène chez la variation génétique existante soient contestées (Goodsman et al., 2014). En effet, outre les expériences faites avec des bactéries par Hallatschek and Nelson (2009), peu de preuves ont été trouvées chez des populations naturelles. Les découvertes de Graciá et al. (2013) en sont une rare exception. Ces chercheurs ont révélé que plusieurs allèles neutres rares chez les populations de tortues nord-africaines sont devenus très communs durant leur colonisation des côtes espagnoles. L'étude de Graciá et al. (2013) et de plusieurs autres (Gehara, Summers, & Brown, 2013; Moreau et al., 2011; Sousa, Peischl, & Excoffier, 2014) indiquent que le surf d'allèles est probablement bel et bien présent dans les populations naturelles.

Le surf d'allèles se déroule comme suit. D'abord, dans les sous-populations qui se forment au front de l'expansion, les allèles sont aléatoirement choisis selon la dérive génétique pour devenir surreprésentés par rapport à la génération précédente (Excoffier & Ray, 2008). Ensuite, comme les groupes d'individus qui renferment ces fréquences alléliques anormalement grandes se dispersent et fondent à leur tour de nouvelles sous-populations au front en progression, il pourrait y survenir un deuxième processus de sélection aléatoire de ces mêmes allèles, causé encore une fois par l'effet fondateur (Slatkin & Excoffier, 2012). Ainsi, simplement déclenché par un processus aléatoire, un allèle qui était initialement rare dans la population (ou une mutation qui surgit durant une expansion) peut devenir largement répandu dans les sous-populations du front de l'expansion (Excoffier et al., 2009). Néanmoins, la probabilité de surf d'un allèle est directement proportionnelle à sa fréquence initiale (Excoffier & Ray, 2008) et sa proximité du front (Burton & Travis, 2008c). Ce phénomène d'effets fondateurs successifs induit donc une baisse graduelle au front de l'expansion de la diversité génétique (Excoffier & Ray, 2008; Peter & Slatkin, 2013) et de l'hétérozygotie (Austerlitz, Jung-Muller, Godelle, & Gouyon, 1997). De plus, comme les fréquences alléliques au front ont une plus grande variance, les allèles sont davantage susceptibles de se fixer ou de s'éteindre localement (Peter & Slatkin, 2015). Bref, le surf d'allèles, si prouvé être un processus génétique communément à l'œuvre dans les populations naturelles, pourrait ainsi avoir un impact important sur le déroulement d'une expansion, en modifiant considérablement la balance des allèles (Excoffier & Ray, 2008).

5.1.3.c. Impacts de la dispersion

Un autre facteur à considérer durant une expansion de territoire est la capacité de dispersion individuelle. Celle-ci influence la dynamique de la population et la structure génétique, liant les processus micro-évolutifs aux patrons observés à l'échelle du paysage (Larroque et al., 2016). La dispersion est un processus complexe qui est étroitement relié au flux génique (Clobert, Le Galliard, Cote, Meylan, & Massot, 2009).

Le flux génique a plusieurs répercussions sur la structure génétique d'une population, mais son action principale est d'introduire dans le bassin génétique d'un groupe d'individus du matériel génétique issu d'un autre groupe d'individus (Mallet, 2001). Théoriquement, un flux génique venant du cœur de la population aura tendance à atténuer, voire même éliminer certaines variantes génétiques qui ont émergées indépendamment en périphérie, soit à cause de la sélection ou encore la dérive génétique (Endler, 1973). Ce flux génique qualifié de « maladaptatif » empêche les sous-populations en périphérie de bien s'adapter aux conditions locales et de repousser les limites de l'aire de répartition (Polechová & Barton, 2015). D'un autre côté, il est attendu que le flux génique soit également responsable d'un influx bénéfique de diversité génétique dans les populations réceptrices dont le matériel génétique est constamment érodé par la dérive génétique, ralentissant ainsi dans une certaine mesure l'action de celle-ci (Bohonak, 1999). Au fil des décennies, les preuves empiriques et théoriques s'accumulent des deux côtés de la médaille : plusieurs études relèvent l'importance d'un flux génique maladaptatif (ex: De Mita et al., 2013; Débarre, Ronce, & Gandon, 2013; Fitzpatrick, Gerberich, Kronenberger, Angeloni, & Funk, 2015; Rolshausen et al., 2015; Roulin et al., 2015), alors que d'autres découvrent un effet plutôt bénéfique ou bien moins nuisible qu'attendu (ex: Bell & Gonzalez, 2011; Dlugosch & Parker, 2008; Fitzpatrick et al., 2015; Schweizer et al., 2016; Tigano & Friesen, 2016; Volis, Ormanbekova, & Shulgina, 2016). Cet antagonisme vient donc complexifier l'étude de la structure génétique et de la distribution d'une espèce (Kirkpatrick & Barton, 1997). Ces processus génétiques surviennent tout au long d'une expansion de territoire et seront grandement influencés par la capacité de dispersion des individus (van Strien, Holderegger, & Van Heck, 2014).

Avec une capacité de dispersion réduite, la progression d'une expansion est susceptible d'occasionner une forte structure génétique sur le territoire, c'est-à-dire une différenciation génétique marquée entre les sous-populations réparties dans l'espace (Cahill & Levinton, 2016).

Ceci s'explique par l'absence de flux génique entre sous-populations éloignées : cet échange de gènes s'effectue indirectement « par étapes » d'une sous-population à une autre, d'une génération à la suivante. Dans un tel scénario, le front de l'expansion plus isolé est prémuni d'un potentiel flux génique maladaptatif venant du cœur, mais il est également privé d'un potentiel influx bénéfique de diversité génétique (Polechová & Barton, 2015). Un tel flux génique limité permet le surf de certains allèles au front de l'expansion (Excoffier & Ray, 2008). De plus, la formation de secteurs spatiaux plus isolés des autres au sein du territoire colonisé est possible avec un faible flux de gènes, ce qui conduit à l'établissement de sous-populations différenciées où il y a dominance de certains génotypes, ce qui occasionne une structure génétique particulière (« sectoring »; Goodsmann et al., 2014).

La dispersion à longue distance, de son côté, est un mécanisme démographique moins bien compris car relativement rare et difficile à observer et quantifier (Nathan, Perry, Cronin, Strand, & Cain, 2003). La dispersion à longue distance, même lorsqu'elle intervient rarement, influence davantage le déroulement génétique et spatial de l'expansion d'une population que les événements de faible dispersion (Neubert & Caswell, 2000). Le principal impact d'un tel mode de dispersion concerne l'échange de matériel génétique beaucoup plus immédiat entre les individus du cœur et ceux de la périphérie, distribuant ainsi les allèles plus ou moins également à travers le territoire (Bialozyt, Ziegenhagen, & Petit, 2006; Peter & Slatkin, 2015). Ceci a tendance à homogénéiser la variation génétique et peut prévenir la fixation d'allèles dans les groupes d'individus isolés du front (De Mita et al., 2013).

5.1.3.d. Effet du temps

Le temps écoulé depuis le début d'une expansion de territoire est déterminant pour la structure génétique puisqu'il reflète le nombre de générations intervenues dans les processus génétiques durant l'expansion. Bien entendu, un échantillon temporel précis aura des particularités génétiques différentes selon l'espèce à l'étude et les conditions démographiques précises.

Une expansion de territoire s'arrête lorsque son front rencontre des conditions biotiques ou abiotiques qui changent de manière trop abrupte pour que la population puisse graduellement s'y adapter (Kirkpatrick & Barton, 1997). À ce moment, les génotypes du cœur de la population pourront définitivement rattraper ceux du front de l'expansion, de manière

directe ou indirecte via le flux génique. Entre eux, il survient alors un mélange du matériel génétique qui aura tendance à homogénéiser la structure génétique sur l'ensemble du territoire (Mallet, 2001) et atténuer la signature génétique de l'expansion (Lotterhos & Whitlock, 2015; Slatkin & Excoffier, 2012). Avec cette stabilisation démographique, il y a également une baisse majeure de l'influence de la dérive génétique et une hausse du poids de la sélection naturelle (Travis et al., 2007). Ainsi, les allèles délétères qui ont proliféré grâce au surf commencent peu à peu à être éliminés de la population : si l'allèle est récessif, il persistera plus longtemps (Slatkin & Excoffier, 2012).

5.2. Cadre théorique

Mon projet de recherche s'inscrit dans le cadre théorique de la génétique et de la génomique du paysage. Je définis ici ces disciplines et discute de la problématique reliée avec les expansions du territoire.

5.2.1. Génétique du paysage

Tous ces processus génétiques et démographiques qui viennent d'être élaborés peuvent être mesurés et quantifiés à l'aide des données génétiques (Fabian et al., 2012; Manel, Schwartz, Luikart, & Taberlet, 2003; Rohlf & Schnell, 1971; Slatkin & Barton, 1989; Wright, 1951). Étudier ces phénomènes devient particulièrement intéressant lorsqu'ils sont placés dans un contexte spatial, en tenant compte de l'effet de l'hétérogénéité du paysage sur ces processus génétiques (Manel & Holderegger, 2013). La génétique du paysage, une ramification de l'écologie du paysage qui a émergé il y a une dizaine d'années avec les progrès techniques en génétique moléculaire (Manel et al., 2003), s'intéresse à ces questions et se concentre sur le rôle de la structure du paysage dans les processus démographiques et de structuration génétique neutre (Rolf Holderegger, Kamm, & Gugerli, 2006). L'attrait de cette branche de l'écologie spatiale vient de sa capacité d'adresser des questions d'intérêt général comme l'impact de variables environnementales sur le mouvement et l'échange de gènes des organismes, ce qui facilite la conservation d'espèces menacées et éclaire l'étude des processus évolutifs (Richardson, Brady, Wang, & Spear, 2016).

5.2.2. Génomique du paysage

La génomique du paysage se concentre sur l'identification de gènes/loci potentiellement adaptatifs et étudie les impacts des caractéristiques du paysage sur cette variation génétique adaptative (Manel & Holderegger, 2013). Elle y arrive grâce à l'accessibilité toujours plus grande

à de gigantesques jeux de données génétiques, qui lui permet d'analyser simultanément des milliers de marqueurs génétiques (Schwartz et al., 2010). L'utilisation de scans de génomes entiers permet de relever des traces de la sélection naturelle sans requérir des connaissances préalables sur les déterminants génétiques des traits phénotypiques (Siol, Wright, & Barrett, 2010). La génomique du paysage est donc très appropriée pour aborder des questions concernant l'évolution et l'adaptation de populations et d'espèces non modèles qui font face aux changements rapides des conditions environnementales (Stéphane Joost et al., 2013; Landguth & Balkenhol, 2012), avec souvent un focus sur la conservation (Manel & Holderegger, 2013).

La génomique du paysage est complémentaire aux techniques non-spatiales déjà existantes pour identifier les loci potentiellement sous sélection, tels l'analyse QTL (« *quantitative trait loci* »), la méthode du gène candidat (« *candidate gene mapping* »), le « *transcription profiling* » (R. Holderegger et al., 2008) ou encore le discernement de régions génomiques démontrant des signes de sélection (Song, Downes, Parker, Oakeshott, & Robin, 2015). D'autres approches ont été développées à partir de ces travaux : par exemple, dans leur étude, Janes et al. (2014) ont utilisé le programme LOSITAN, qui se base sur une méthode de « *F_{ST}-outlier* » pour détecter des adaptations moléculaires, alors que Pariset, Joost, Ajmone Marsan, and Valentini (2009) emploient une autre méthode (programme « *FDIST2* ») basée sur une mesure de F_{ST} pour identifier des loci adaptatifs. Ces techniques similaires relèvent les loci avec des différenciations génétiques inter-populationnelles qui sont anormales par rapport à un modèle nul de gènes neutres (Beaumont & Balding, 2004). Un autre groupe de méthodes tente plutôt d'identifier les loci qui sont significativement corrélés avec une ou plusieurs variables environnementales (Rellstab, Gugerli, Eckert, Hancock, & Holderegger, 2015). C'est le cas notamment des méthodes LFMM (« *latent factor mixed models* »; Frichot, Schoville, Bouchard, & François, 2013), de l'utilisation d'une matrice de covariance par BAYENV (Coop, Witonsky, Rienzo, & Pritchard, 2010) et de la méthode SAM (« *spatial analysis method* »), qui se sert de données de systèmes d'information géographique (S. Joost et al., 2007; S. Joost, Kalbermatten, & Bonin, 2008).

5.2.3. Limites et problématique

La génétique et la génomique du paysage connaissent depuis une décennie une véritable effervescence. Les études touchent à un vaste domaine d'applications : de l'impact des structures du paysage sur la connectivité et la structure génétique (ex: Coulon et al., 2004;

Coulon et al., 2006; Spear, Peterson, Matocq, & Storfer, 2005), à la recherche de signatures génétiques précises (ex: Oyler-McCance et al., 2016; Rózsa, Strand, Montadert, Kozma, & Höglund, 2016), en passant par l'identification de gènes potentiellement adaptatifs (ex: Schweizer et al., 2016; Swaegers et al., 2015; Zenni & Hoban, 2015). Dans ces derniers cas, dépister les gènes influencés par la sélection naturelle comporte certaines problématiques qui influencent négativement les résultats d'analyse (Haasl & Payseur, 2016).

En effet, les méthodes de corrélation environnementale couramment employées en génomique du paysage font parfois fi de quelques mécanismes démographiques et génétiques importants. D'abord, dans ce type de méthodes, il y a supposition que la sélection a pu agir suffisamment longtemps pour établir une relation observable entre la distribution des allèles et une variable environnementale (Stéphane Joost et al., 2013). Cette hypothèse n'est pas valable pour beaucoup de populations stables, mais encore moins pour les populations qui viennent d'achever une expansion de territoire ou qui sont toujours en pleine expansion (Epps & Keyghobadi, 2015). D'ailleurs, en ce qui concerne l'adaptation aux changements environnementaux en cours, ceux-ci peuvent conduire à certaines adaptations après quelques générations, mais seulement si la force de sélection est suffisamment forte (Kawecki & Ebert, 2004). Ensuite, une difficulté majeure vient de la forte structuration de la population, qui induit certaines « duperies » dans le matériel génétique. Comme la capacité de dispersion détermine le niveau de flux génique entre sous-populations plus ou moins éloignées, une faible dispersion réduira l'échange génétique entre les extrémités d'un territoire, augmentant ainsi la différenciation génétique (Nei, 1973c). De la sorte, pour un gène neutre, l'action combinée du flux génique limité et de la dérive génétique locale pourra créer un motif d'isolation par distance qui se traduira en un gradient de fréquences alléliques sur le territoire (Vasemägi, 2006). Ce gradient, qui est similaire à ceux attendus avec la variation génétique adaptative, peut donc être généré en l'absence de variation géographique dans l'intensité de la sélection (Slatkin & Excoffier, 2012). Une méthode de corrélation environnementale qui ignore ainsi les motifs d'isolation par distance pourra donc relever des corrélations fallacieuses (c.-à-d. des faux-positifs), en identifiant des gènes neutres comme étant potentiellement adaptatifs (François et al., 2016; Parisod & Holderegger, 2012).

Néanmoins, ces dernières années ont connu un grand engouement pour le développement de méthodes sophistiquées qui sauraient régler les diverses problématiques associées à l'identification de gènes potentiellement adaptatifs (Jensen, Foll, & Bernatchez, 2016). Comme le défi est en effet de distinguer les signaux adaptatifs à travers tout le bruit démographique du matériel génétique (Flatt, 2016), il devient nécessaire de prendre en compte ce motif d'isolation par distance (Meirmans, 2012). Ainsi, plusieurs méthodes d'association environnementale préexistantes furent raffinées pour tenir compte de cette problématique. Par exemple, Günther and Coop (2012) ont amélioré l'approche bayésienne de BAYENV, qui fut ensuite retravaillée et perfectionnée par Gautier (2015) pour améliorer sa capacité d'identifier correctement les gènes adaptatifs. Par contre, selon les méthodes, certains facteurs confondants, comme le déséquilibre de liaison, le plan d'échantillonnage ou encore la proximité génétique entre les individus, ne sont pas considérées (François et al., 2016). C'est pourquoi certains chercheurs préconisent l'utilisation et le développement d'approches multivariées, telles les ordinations, pour identifier les gènes potentiellement adaptatifs tout en contrôlant pour la structure de la population et autres facteurs confondants (Duforet-Frebourg, Luu, Laval, Bazin, & Blum, 2015; Forester et al., 2016; Jombart, Pontier, & Dufour, 2009). G.-B. Chen, Lee, Zhu, Benyamin, and Robinson (2016) ont d'ailleurs développé une approche d'association environnementale basée sur la décomposition des données par des vecteurs propres (« *EigenGWAS* »). Grâce à tous ces développements et améliorations, les méthodes d'association environnementale ont un pouvoir de détection largement supérieur à celui des méthodes basées sur les mesures de F_{ST} (De Mita et al., 2013). Cependant, tous les processus génétiques caractéristiques d'une expansion de territoire, comme la dérive génétique et le surf d'allèles, sont susceptibles de confondre les résultats des méthodes employées en génomique du paysage et de gonfler le taux de faux-positif (Frichot et al., 2015). C'est pourquoi le but de ce projet est de mieux comprendre dans quelles conditions nous rencontrons un plus grand nombre d'allèles neutres qui démontrent des patrons spatiaux semblables à ceux produits par des allèles adaptatifs durant une expansion de territoire.

5.3. Le dendroctone du pin

Mon projet de recherche explore la structuration génétique d'une population en expansion de territoire en étudiant le cas du dendroctone du pin. Cette section résume la biologie et l'épidémie en cours de cet insecte ravageur.

5.3.1. Résumé de la biologie

Dans l'ouest du Canada, les écosystèmes sont présentement grandement affectés par l'expansion territoriale du dendroctone du pin (*Dendroctonus ponderosae*), un insecte ravageur indigène des forêts de pins de l'ouest d'Amérique du nord (Cudmore, Björklund, Carroll, & Lindgren, 2010). Ce coléoptère est de la famille des Curculionidae et de la sous-famille des Scolytinae, qui comporte environ 250 genres et plus de 6000 espèces de petits insectes xylophages (c.-à-d. qui se nourrit de matière ligneuse) et qui était autrefois considérée comme une famille distincte (Jordal, 2014). Le genre *Dendroctonus* comprend 16 espèces en Amérique, répartis du Nicaragua jusqu'en Arctique, et une espèce supplémentaire présente sur le territoire européen et asiatique (S. Wood, 1982) qui ont toutes des associations millénaires et étroites avec les conifères (Sequeira, Normark, & Farrell, 2000). Le dendroctone du pin est une espèce majoritairement univoltine, c'est-à-dire qu'une seule génération est produite annuellement et que les individus ne vivent en général pas plus d'un an (Logan & Bentz, 1999). Au Canada, les populations infestent principalement le pin tordu (*Pinus contorta*), mais d'autres essences conifériennes sont également susceptibles d'être affectées (S. L. Wood, 1963). La femelle adulte, en été, sélectionne un hôte potentiel et produit des phéromones d'agrégation qui attirent des centaines de dendroctones mâles et femelles (Pitman & Vité, 1969). Les femelles creusent alors des trouées dans l'écorce de l'arbre et répandent dans les tissus de l'hôte des spores de diverses espèces de champignons pathogènes qui interrompent le transport et la production de la sève (Addison, Powell, Six, Moore, & Bentz, 2013). L'action combinée de l'attaque massive des dendroctones et du champignon aura raison de la défense de l'hôte, qui finira par mourir (Lieutier, Yart, & Salle, 2009). Les femelles construisent des galeries verticales sous l'écorce de l'arbre et pondent des œufs tout au long de celles-ci (Amman & Cole, 1983). Les larves émergées se nourrissent alors du phloème de l'hôte durant plusieurs mois (Amman & Cole, 1983). Leur développement et leur survie hivernale sont grandement influencés par la température : le cycle de vie du dendroctone du pin est ainsi assez sensible aux conditions climatiques (Preisler, Hicke, Ager, & Hayes, 2012). L'émergence des nouveaux adultes à l'été suivant démarre un nouveau cycle et enclenche alors l'attaque de nouveaux hôtes environnants (Les Safranyik & Wilson, 2006).

5.3.2. Capacité de dispersion

La capacité de dispersion des dendroctones du pin adultes est un paramètre relativement méconnu qui fait présentement l'objet de nombreuses études (Evenden, Whitehouse, & Sykes, 2014). Généralement, les adultes se dispersent à une très faible distance, trouvant un hôte adéquat dans le même peuplement que leur lieu d'émergence (H. Chen & Walton, 2011). Toutefois, des individus sont parfois emportés par les grands courants aériens surplombant la canopée et peuvent ainsi parcourir des dizaines de kilomètres avant d'atterrir dans un nouvel habitat (L. Safranyik, Linton, Silversides, & McMullen, 1992). Cette dispersion à longue distance difficilement prédictible est reconnue comme étant une composante importante des épidémies de dendroctones, composante qui complique beaucoup l'élaboration de modèles prédictifs d'expansion de territoire (L. Safranyik et al., 2010).

5.3.3. Épidémies

Originaires de la moitié méridionale de la Colombie Britannique, les populations épidémiques canadiennes de dendroctones se propagent rapidement et continuellement depuis la fin des années 90 vers des habitats plus au nord et à l'est (Janes et al., 2014). Bien que le système épidémiologique soit historiquement bien étudié et validé, l'épidémie qui secoue actuellement l'ouest canadien est d'une ampleur et d'une nouveauté bien unique. En effet, les populations de dendroctones connaissent des expansions de territoire qui le font déborder de son aire de répartition historique, franchissant les rocheuses et s'établissant dans de nouveaux sites au nord de la Colombie-Britannique (de la Giroday, Carroll, & Aukema, 2012). Autrefois confinées à l'ouest des Rocheuses, les populations sont désormais présentes dans le nord albertain (Janes et al., 2014).

Plusieurs hypothèses ont été émises pour tenter d'expliquer la présence d'un tel phénomène. D'abord, la structure du paysage semble expliquer une partie des patrons observés. Par exemple, de la Giroday, Carroll, Lindgren, and Aukema (2011) ont identifié certaines structures fréquemment colonisées par les dendroctones dispersants, tels les canyons, les vallées et les basses collines. Aussi, le dendroctone répond positivement à la fragmentation des forêts et à l'augmentation de bordures créées par l'industrie forestière (H. Chen, Ott, Wang, & Ebata, 2014). En effet, les conditions microclimatiques générées par les bordures forestières et l'effet de lisière sont favorables à l'infestation réussie d'un peuplement (Kautz, Schopf, & Ohser, 2013). Plus spécifiquement, les habitats fragmentés sont davantage exploités par les

populations de dendroctones à faible densité d'individus (c.-à-d. durant une phase endémique) : la communication par les phéromones et les déplacements sont alors facilités, les arbres hôtes sont plus fragiles, l'hiver est légèrement moins rigoureux (car plus d'ensoleillement) et l'émergence au printemps plus précoce (Boone, White, Wulder, Robertson, & Nelson, 2013). Lors d'une phase épidémique, où les individus se comptent par millions, les populations exploitent notamment les territoires forestiers continus étant donné la facilité de surmonter les défenses des hôtes et la plus grande accessibilité immédiate de ressources (Boone et al., 2013). Il est important de noter, également, que la dispersion n'a pas (encore) été catalysée par le transport anthropique du bois de chauffage contaminé (de la Giroday et al., 2012).

Une autre perspective se penche sur les caractéristiques génétiques de cet insecte pour expliquer son expansion. Le dendroctone du pin, historiquement, occupait une vaste étendue latitudinale (du Mexique en Colombie-Britannique) et était donc déjà adapté à une gamme de températures importante (Samarasekera et al., 2012), ce qui le prédisposait probablement à exploiter les variations encourues par les changements climatiques (L. Safranyik et al., 2010). De plus, Janes et al. (2014) suggèrent que les populations en expansion ont pu ajuster leurs fonctions métaboliques et cellulaires à l'échelle du génome afin de mieux supporter les climats plus froids, mais aussi pour faciliter les dispersions à longue distance. En bref, on considère qu'un amalgame de changements climatiques, de perturbations anthropiques et de conditions climatiques et génétiques favorables est susceptible d'avoir permis l'atteinte de seuils critiques pour l'expansion démographique et territoriale du dendroctone du pin (Raffa et al., 2008).

5.3.4. Conséquences de l'épidémie en cours

Bien entendu, une épidémie de cette ampleur a de grandes conséquences sur les écosystèmes affectés et l'économie canadienne. En 2010, plus de 15 millions d'hectares de forêt avaient été affectés en Colombie-Britannique seulement depuis le début de l'épidémie (L. Safranyik et al., 2010). Le pin tordu latifolié (*Pinus contorta* var. *latifolia*), la nourriture de prédilection du dendroctone, représente dans l'ouest canadien l'espèce de pin la plus économiquement viable et profitable (Pedersen, 2003). L'acquisition récente d'un nouvel hôte de l'est, le pin gris (*Pinus banksiana*), menace désormais les écosystèmes boréaux qui s'étendent dans le reste du pays (Cullingham et al., 2011). Certaines provinces présentement non affectées par l'épidémie, comme la Saskatchewan, le Manitoba et l'Ontario, investissent des fonds considérables et participent activement dans la recherche pour mieux atténuer et évaluer les

risques encourus (Cooke, 2013). Avec la décimation de grandes étendues forestières, l'émission de carbone est également un souci majeur, le cycle de carbone étant sensiblement altéré depuis le début de l'épidémie (Kurz, Dymond, et al., 2008).

5.3.5. *Structure génétique du dendroctone du pin*

Historiquement il y avait en Colombie-Britannique deux groupes génétiques et géographiques significatifs : le groupe du nord et le groupe du sud (Samarasekera et al., 2012). Il a été déterminé que les populations nordiques ne présentaient aucun motif d'isolation par distance (c.-à-d. un matériel génétique relativement homogène) et que celles-ci ont été le point de départ des dispersions de populations qui ont franchi les Rocheuses (Samarasekera et al., 2012).

6. Hypothèses

Dans le cadre de ce projet de maîtrise, je m'attends à rencontrer avec mon modèle plusieurs phénomènes génétiques particuliers aux expansions de territoire. Par exemple, j'anticipe la présence d'importants effets fondateurs qui induiront parfois du surf d'allèles. Ceci transparaîtra dans mes résultats où seront observés certains allèles neutres dont la fréquence allélique affichera une forte corrélation avec la distance géographique. Je m'attends à ce qu'il y ait plus de faux-positifs lorsque l'expansion de territoire (1) est dirigée par une faible capacité de dispersion, (2) sur un paysage à faible gradient de sélection et que (3) une population échantillonnée tôt durant l'expansion de territoire conduira à un plus haut taux de faux-positifs.

Chapitre 2: How range expansion can confound landscape genomic approaches to identify loci under selection

Paul Mayrand¹, Élise Filotas², Patrick M. A. James¹

¹Université de Montréal, Département de sciences biologiques, 90 av. Vincent-d'Indy, Montréal, Qc, Canada, H2V 2S9

²TÉLUQ (Université du Québec), Département Science et Technologie, 5800 rue Saint-Denis, Montréal, Qc, Canada, H2S 3L5

In preparation for submission to: Molecular Ecology

Abstract

The identification of adaptive loci is important to better understand evolutionary potential and invasiveness of species undergoing range expansion. However, in irruptive expanding species such as the mountain pine beetle (MPB), demographic processes such as allele surfing and IBD can result in spatial patterns of neutral genetic variation that mimic those that result from adaptive processes. This phenomenon inflates the false discovery rate of adaptive loci and confounds approaches to identify loci under selection. The influence of these demographic and genetic processes on the extent of this phenomenon in expanding populations is still mostly unknown. In this paper, we investigate how different demographic conditions associated with range expansion influence the rate of neutral loci mimicking the spatial patterns of adaptive loci, using the mountain pine beetle outbreak system as a case study. We simulated the demographic and population genetic dynamics of mountain pine beetle populations undergoing range expansion using the spatially explicit, individual-based genetic model CDmetaPOP. We examined the consequences on the false discovery rate of three factors: 1) species dispersal capacity; 2) timing of sampling during the course of the expansion; and 3) the strength of selection on adaptive reference loci. We found that a combination of weak dispersal capacity, weak selection, and early sampling during expansion results in the highest rate of false positive, while strong dispersal was responsible for lower rates of neutral loci mimicking adaptive ones. Used under these conditions of range expansion, landscape genomics models risk elevated false discovery rates of adaptive loci and must be interpreted cautiously. Complex demography in the MPB system and other irruptive species makes identification of adaptive loci challenging. We clearly showed that there is a need for further method development to include these directional demographic processes in the field of landscape genomics.

Keywords: mountain pine beetle, range expansion, landscape genomics, CDmetaPOP, adaptive genetic variation, neutral genetic variation, dispersal, selection

Introduction

Range expansion is a complex demographic phenomenon that can be the result of multiple processes including directional species adaptation to environmental changes (Hassall et al., 2009), species introduction in new habitat (e.g., species invasions) (Brown et al., 2015; Lombaert et al., 2014), or population irruptions (Raffa, Powell, & Townsend, 2013). Range expansions have the potential to be damaging to ecosystem integrity and human economy (Pejchar & Mooney, 2009). For instance, many native irruptive species are currently exhibiting spatial outbreak patterns that exceed the spatial range of their coevolved hosts and associated ecological communities, disturbing the ecological structure (i.e., altering resources availability and ecological interactions between species) in novel habitats (Logan, Régnière, & Powell, 2003; Raffa et al., 2013).

Range expansion consequences for population genetic structure confound equilibrium models (Excoffier et al., 2009). For example, for individuals able to disperse great distances, a population can be characterized by heterogeneous pockets of weak genetic diversity persisting throughout the expansion (Bialozyt et al., 2006). In other cases, successive founder effects can create genetic clines in neutral and adaptive loci along the expansion axis (Peter & Slatkin, 2015). It is thus important to improve our understanding of the underlying demographic and evolutionary mechanisms involved in successful range expansion in order to better manage, predict or prevent the spread of these species.

One way to improve our understanding of the spatial-temporal eco-evolutionary dynamics of range expansions is to identify spatial variation in adaptive loci linked to phenotypic traits important for range expansion and colonization (i.e., landscape genomics; R. Holderegger et al., 2008). Characterization of these adaptive spatial genetic patterns can provide insight into the processes driving range expansion and are needed to guide management decisions that aim to slow the expansion or mitigate its economic and ecological consequences.

Many landscape genomic tools currently exist to identify geno-environmental correlations, including BAYENV (Coop et al., 2010), LFMM (Frichot et al., 2013) and SAM (S. Joost et al., 2007). Recently, additional methods have been developed that account for the confounding effects of neutral population genetic structure (Frichot et al., 2013), sampling design (Günther & Coop, 2012), and linkage disequilibrium (François et al., 2016). However, to our knowledge, the

potentially confounding effects of range expansion and allele surfing on identifying loci under selection using landscape genomics is still largely unknown and none of these methods fully takes this problem into consideration.

Range expansion can create spatial patterns in neutral genetic variation (i.e., in allele frequencies) that can mimic those expected due to adaptive processes (De Villemereuil, Frichot, Bazin, François, & Gaggiotti, 2014). This interaction between complex demography and population genetic processes increases the probability of identifying neutral loci as adaptive (i.e., inflated FDR; Storey & Tibshirani, 2003). Two phenomena contribute to inflation of the false discovery rate of putatively adaptive loci. First, neutral population structure as a result of isolation by distance (IBD) or isolation by resistance (IBR) can cause differentiation in neutral allele frequencies across the landscape (Nei, 1973a) and may include neutral clines (Vasemägi, 2006). Second, allele surfing during range expansion can cause neutral loci to behave similarly to adaptive loci. Allele surfing occurs when alleles of either neutral or adaptive loci reach higher frequencies than expected near the front of the range expansion as a result of serial founder effects (Klopfstein et al., 2006; Slatkin & Excoffier, 2012).

Despite our awareness of these issues, there is still a lack of understanding of how allele surfing changes with demographic contexts and through time and how the resulting patterns affect current approaches to identifying loci under selection. In this paper, we used the mountain pine beetle (*Dendroctonus ponderosae*) (MPB) system as a model of range expansion. We employed the spatially explicit, individual-based model CDmetaPOP (Landguth et al., 2016) to simulate range expansions by parametrizing our model with MPB outbreak and life history traits. We then investigated with simulated data the conditions under which allele frequencies of neutral loci are more likely to exhibit spatial patterns (i.e., clines) similar to those expected of loci undergoing selection. Specifically, we quantified how the probability of incorrectly identifying a neutral locus as adaptive varies as a function of: 1) organism dispersal capacity; 2) timing of sampling during range expansion; and 3) strength of selection for a reference adaptive locus.

Mountain pine beetle range expansion

The mountain pine beetle is an irruptive forest pest in western North America that attacks and kills pine species (Les Safranyik & Wilson, 2006). An ongoing MPB outbreak (Raffa et al., 2008) continues to expand along two geographically distinct axes: one in northern British Columbia,

moving towards Alaska and the Yukon Territory, and one in Alberta, moving east towards Saskatchewan and eastern Canadian provinces (L. Safranyik et al., 2010). By 2010, this outbreak has killed more than 15 million hectares of pine forest in Canada and ongoing range expansion poses a number of economical and ecological risks (Cullingham et al., 2011). Indeed, MPB's main host (i.e., lodgepole pine; *Pinus contorta*) is the most economically important pine species for the forest industry in western Canada (Pedersen, 2003). Moreover, with the death of large forest stands, the global carbon cycle of western Canada has been profoundly altered in recent years, leading the country toward lower carbon capture (Kurz, Stinson, Rampley, Dymond, & Neilson, 2008).

Mountain pine beetle outbreak dynamics are complex and shaped by the interactions among multiple spatial and temporal factors including landscape topography (de la Giroday et al., 2011), forest fragmentation (Kautz et al., 2013), adult beetle dispersal (L. Safranyik et al., 1992), connectivity among patches of mature lodgepole pine forest (Boone et al., 2013), climate (James, Coltman, Murray, Hamelin, & Sperling, 2011), and close association with fungal symbionts (Roe, James, Rice, Cooke, & Sperling, 2011). Range expansion may have been influenced by local adaptation to new habitat environmental conditions, as suggested by Janes et al. (2014). Indeed, the current MPB outbreak in western Canada has undergone rapid range expansion (L. Safranyik et al., 2010) and the expansion has allowed the eastern population to successfully colonize a new host species (i.e., jack pine, *Pinus banksiana*; Cullingham et al., 2011). However, the contribution of adaptation in the current outbreak and range expansion of MPB is still largely unknown.

Many efforts are underway to identify loci associated with adaptive traits such as cold tolerance, flight capacity, and terpene detoxification, following the groundwork done by Adams et al. (2013), Bentz, Bracewell, Mock, and Pfreder (2011), and Bracewell, Pfreder, Mock, and Bentz (2013). Characterization of the spatial distribution of such adaptive loci will improve understanding of the evolutionary mechanisms driving MPB colonization success and invasive capacity; an essential step to predict future outbreak and expansions and to evaluate the risk to boreal forests, including jack pine forest stands, east of the current outbreak extent (Cullingham, James, Cooke, & Coltman, 2012). While landscape genomics methods are compelling tools to reach these goals, uncertainty remains regarding the expected production of a high false

discovery rate of putatively adaptive loci in the complex demographic situation of the current MPB outbreak.

Methods

Modelling framework

The development of spatial genetic structure during range expansion was simulated using CDMetaPOP (Landguth et al., 2016). This model simulates the spatial and temporal demographic dynamics of a population of diploid individuals that represents a simplified version of the MPB system. Each of these individuals has a genotype of 100 neutral loci and one adaptive locus. The landscape was modelled as a homogeneous rectangular grid of 20 by 80 cells with absorbing boundary conditions. Mirroring the geographic extent of MPB range expansion from northwestern British Columbia to northern Alberta within 10 years of the current outbreak (Janes et al., 2014), the simulated landscape can be interpreted as an area of 160 by 650 km with cells of 8,125 by 8,125 km (6600 ha). The carrying capacity of each of those cells was 30 individuals, which translates into a maximum possible population of 48 000 individuals when the simulation grid is full. Although this figure is far from the billions of MPB individuals found in an actual affected landscape, it represents a necessary computational compromise that also satisfies the demographic requirements for the genetic processes under study. The spatiotemporal genomic dynamics simulated using CDMetaPop are the result of three main processes: 1) dispersal, 2) reproduction, and 3) selection.

Dispersal

Dispersal was simulated using a simple dispersal kernel in which the probability of dispersal declines exponentially as a function of the Euclidean distance between cells and which allows long-distance dispersal events:

$$\text{Probability of dispersal} = \beta * 10^{-\alpha * d_{ij}}$$

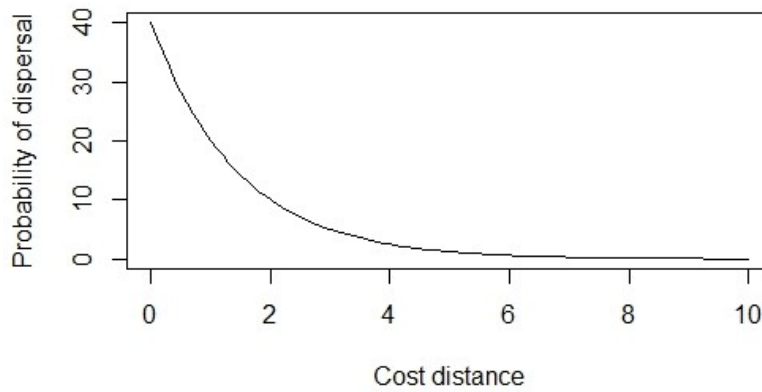
Where d_{ij} is the distance between locations i and j . We set $\beta = 40$ and $\alpha = 0,3$ in order to reach a compromise where most dispersal events are short-distance (i.e., close to the focal cell) while allowing the presence of long-distance dispersal events. To examine the effect of different dispersal capacities on the creation of spatial genetic structure, we varied maximal dispersal distance possible with the truncation of our exponential dispersal kernel (see Figure 1 for more

details). This form of dispersal kernel adequately simulates MPB dispersal as both short (< 2 km according to Robertson, Nelson, Jelinski, Wulder, and Boots (2009)) and long-distance (> 2 km) dispersal events are observed (H. Chen & Walton, 2011).

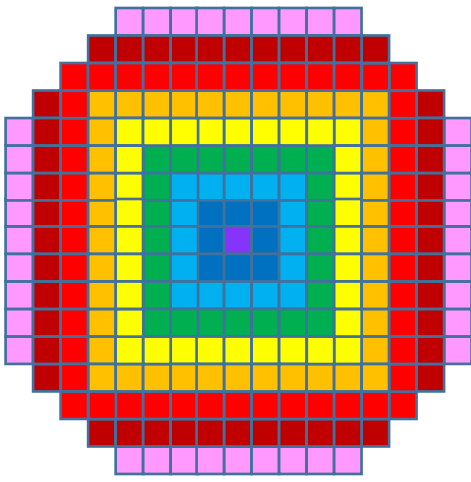
Figure 1: Dispersal model. **A.** Probability of dispersal function and distribution of probabilities, where d is the distance between locations i and j . **B.** Different possible areas of dispersal from a focal cell, marked with distinct colors. **C.** Probability of dispersal in each of the coloured ring areas (leaving from the center), for each chosen maximum dispersal threshold values.

A.

$$\text{Probability of dispersal} = 40 * 10^{-0,3*d_{ij}}$$



B.



C.

Threshold	[0;1]	[1;2]	[2;3]	[3;4,24]
1,99 (Weak)	0,6670	0,3330	0	0
2,99 (Intermediate)	0,5066	0,3627	0,1307	0
8,99 (Strong)	0,3117	0,2732	0,1872	0,1145

[4;5,66]	[5;7,07]	[6;8,49]	[7;8,6]	[8;8,94]
0	0	0	0	0
0	0	0	0	0
0,0605	0,0291	0,0167	0,0056	0,0016

Reproduction

Reproduction was Mendelian and sexual partners were picked from within a same cell after dispersal. As in the MPB system, females produced in average 60 eggs per mating event (Les Safranyik & Wilson, 2006). Also, sex-ratio was set to one male for two females and males could reproduce with more than one female, the simulator pairing every mature female to only one male at each generation.

Selection

To compare neutral loci spatial patterns to an adaptive reference, we simulated one codominant adaptive locus. Selection for this locus was modelled using different geographic selection surfaces represented as a genotype-associated mortality gradient. The bi-allelic adaptive locus had three genotypes: AA , Aa and aa . For each of these genotypes, a table of percentages of mortality was specified, following a south-north gradient across the simulated grid. We created a selection framework such that allele a was selected against in the upper area of the simulation landscape ($aa = 1 - s$, where s is the selection coefficient). The other allele (A) was not affected by selection and was neither advantageous nor deleterious ($AA = 1$). However, the non-selected allele (A) indirectly benefits individuals through a reduced mortality probability (i.e., absence of selection pressure due to allele a). The heterozygote genotype Aa exhibited perfect co-dominance and thus experienced mortality, but half as strongly as the homozygote aa ($Aa = 1 - 0.5s$). This selection framework represents a scenario in which dispersing MPB individuals are leaving a well-suited habitat to a less suited one.

Simulation scenarios

All simulations began with a small population of 120 individuals confined in four adjacent cells in the south part of the territory. A single burn-in generation, under panmictic conditions (i.e. random mating), was simulated at the beginning of every simulation to remove the artefacts of initial random genotype assignment and to establish Hardy-Weinberg equilibrium. The population was then free to expand into the simulation grid. Other simulation parameters are described in Table S.1.

We examined the model sensitivity to three different levels of dispersal (weak, intermediate and strong), three strength of selection (very weak, weak, intermediate, strong, and very strong), and timing of sampling since the beginning of the expansion, resulting in a total of 15 scenarios

(Table I). A weak dispersal capacity corresponds to a maximal dispersal of one cell (i.e., 8,125 km), with at least a third of individuals dispersing from their natal cell (Figure 1). An intermediate dispersal capacity corresponds to a maximal dispersal of two cells (i.e., 16,25 km), with more than 10 % of individuals dispersing two cells away. Finally, strong dispersal capacity corresponds to a maximal dispersal of eight cells (i.e., 65 km); fewer than 1 % of individuals reach this maximal distance.

Five strengths of selection pressure were simulated by varying the slope of the south-north mortality gradient (i.e., very weak, weak, intermediate, strong, and very strong). The five levels of maximum mortality probability at the farthest end of the expansion grid for the homozygous genotype *aa* were 1 %, 5 %, 30 %, 50 %, and 70 %. Homozygote dominant alleles (*AA*) were not affected by the selection gradient, and heterozygotes (*Aa*) experienced a maximal mortality of 0,5 %, 2,5 %, 15 %, 25 %, and 35 % respectively. We replicated each scenario 50 times to capture the range of stochastic variation in the simulated spatial processes (Table I).

Scenario	Dispersal capacity (maximal dispersal)	Strength of selection (maximal mortality rate for <i>aa</i>)	Time to fill the grid (average number of generations and standard deviation)
1	Weak (1 cell)	Very weak (1%)	150 (± 2)
2	Weak (1 cell)	Weak (5%)	151 (± 2)
3	Weak (1 cell)	Intermediate (30%)	167 (± 5)
4	Weak (1 cell)	Strong (50%)	166 (± 4)
5	Weak (1 cell)	Very strong (70%)	165 (± 4)
6	Intermediate (2 cells)	Very weak (1%)	103 (± 2)
7	Intermediate (2 cells)	Weak (5%)	103 (± 2)
8	Intermediate (2 cells)	Intermediate (30%)	112 (± 2)
9	Intermediate (2 cells)	Strong (50%)	112 (± 3)
10	Intermediate (2 cells)	Very strong (70%)	112 (± 2)
11	Strong (8 cells)	Very weak (1%)	53 (± 1)
12	Strong (8 cells)	Weak (5%)	54 (± 1)
13	Strong (8 cells)	Intermediate (30%)	61 (± 3)
14	Strong (8 cells)	Strong (50%)*	60 (± 2)
15	Strong (8 cells)	Very strong (70%)	61 (± 3)

* Due to manipulation mistakes on the data, only 49 replicates were retained for this scenario.

Table I: Summary of the scenarios combinations of parameters: 3 dispersal capacities x 5 strengths of selection = 15 different scenarios. Also, number of time steps needed to complete the range expansion (i.e., when the simulated grid is filled), averaged on all 50 replicates from each scenario.

Identifying outlier loci

To characterise how the false discovery rate of putatively adaptive loci is influenced by dispersal, selection, and timing of sampling, we examined the slopes of linear regressions between allele

frequencies and geographic location for each simulated locus. We summarized the distribution of slopes of the neutral alleles and the slope of the single adaptive locus, which served as a reference for adaptive genetic spatial patterns, at different points in time. The value of neutral allele slopes in response to the experimental treatments and through time, relative to the slope of the known adaptive locus, indicates the conditions under which neutral genetic variation is more likely to exhibit a pattern expected from adaptive genetic variation.

We averaged the slope values for adaptive alleles over all replicates for each combination of dispersal capacity, strength of selection and time. We then counted the number of neutral alleles whose slope was greater than or equal to this reference value. Three-way ANOVA was used to quantify the relative importance of the different experimental factors on the probability of misidentifying neutral loci as potentially adaptive. In this model dispersal capacity, strength of selection, six sampling timings (i.e., generations 25, 50, 75, 100, 125 and 150), and their interactions were used as explanatory variables and the count of neutral “outlier” loci were used as response variable. There were 50 observations for each treatment combination (Table I) reflecting the 50 simulation replicates for each experiment. ANOVA was implemented using R.

Model validation

Prior to launching simulations we validated that the model was producing the patterns expected of neutral loci during range expansion, such as allele surfing. Model behaviour was validated for all different scenarios. We specifically sought to verify that: 1) spatial genetic structure changes with time since the expansion began, 2) greater dispersal results in reduced population genetic structure, and 3) allele surfing during range expansion was resulting in neutral clines along the main axis of expansion (Excoffier & Ray, 2008; Vasemägi, 2006).

To visualize the development of spatial genetic structure through time, we used unconstrained ordination (PCA) on log-transformed neutral allele frequencies from a sample of grid locations ($n = 200$) at six different points in time (i.e., generations 25, 50, 75, 100, 125 and 150). Ordinations were performed at these six generations using data from three replicates (i.e., three simulations) of each scenario, to consider the variability of results issued from stochastic processes. Ordination were undertake using the *rda* function in the *vegan* package in R (Oksanen et al., 2016). Overall population genetic structure was quantified using the unbiased global F_{ST} (Weir & Cockerham, 1984) calculated using the *wc* function in the ‘*HIERFSTAT*’ package

(Goudet & Jombart, 2015) in R. The creation of neutral clines was verified by assessing the relationship between allele frequency and geographic location (latitude) for several randomly sampled neutral loci.

Results

Simulations

Range expansions took on average 160 (± 8) generations to fill the simulated landscape in weak dispersal scenarios, while it took 109 (± 5) generations for an intermediate dispersal capacity, and 58 (± 4) for a strong dispersal capacity (Table I). Time required to fill the simulation grid was lower for simulations under very weak and weak strengths of selection.

Model validation

Our model behaved as expected, producing genetic patterns characteristic of a range expansion (Figure 2). Genetic differentiation increased through time during range expansion in all scenarios, although to a lesser degree with stronger dispersal capacities (Figure 2E-F). When comparing PCAs (Figure 2A-F) from different sampling timings, the genetic structure changes through time; populations far from one another become more genetically distinct. Also, populations have less structure when dispersal capacity is strong: populations in cells far apart can be genetically as similar as in neighbouring cells. Global F_{ST} values (Figure 2A-F) illustrate this increase in spatial genetic structure with time between the middle and the end of the simulation. The pattern of decreasing genetic structure with stronger dispersal capacity is also reflected in the decreasing global F_{ST} values. As expected, strength of selection did not influence neutral genetic structure (results not shown). Finally, neutral clines were frequently identified, across all scenarios and at various timing of sampling during the expansions, indicating IBD pattern and the possible occurrence of allele surfing (Figure 3).

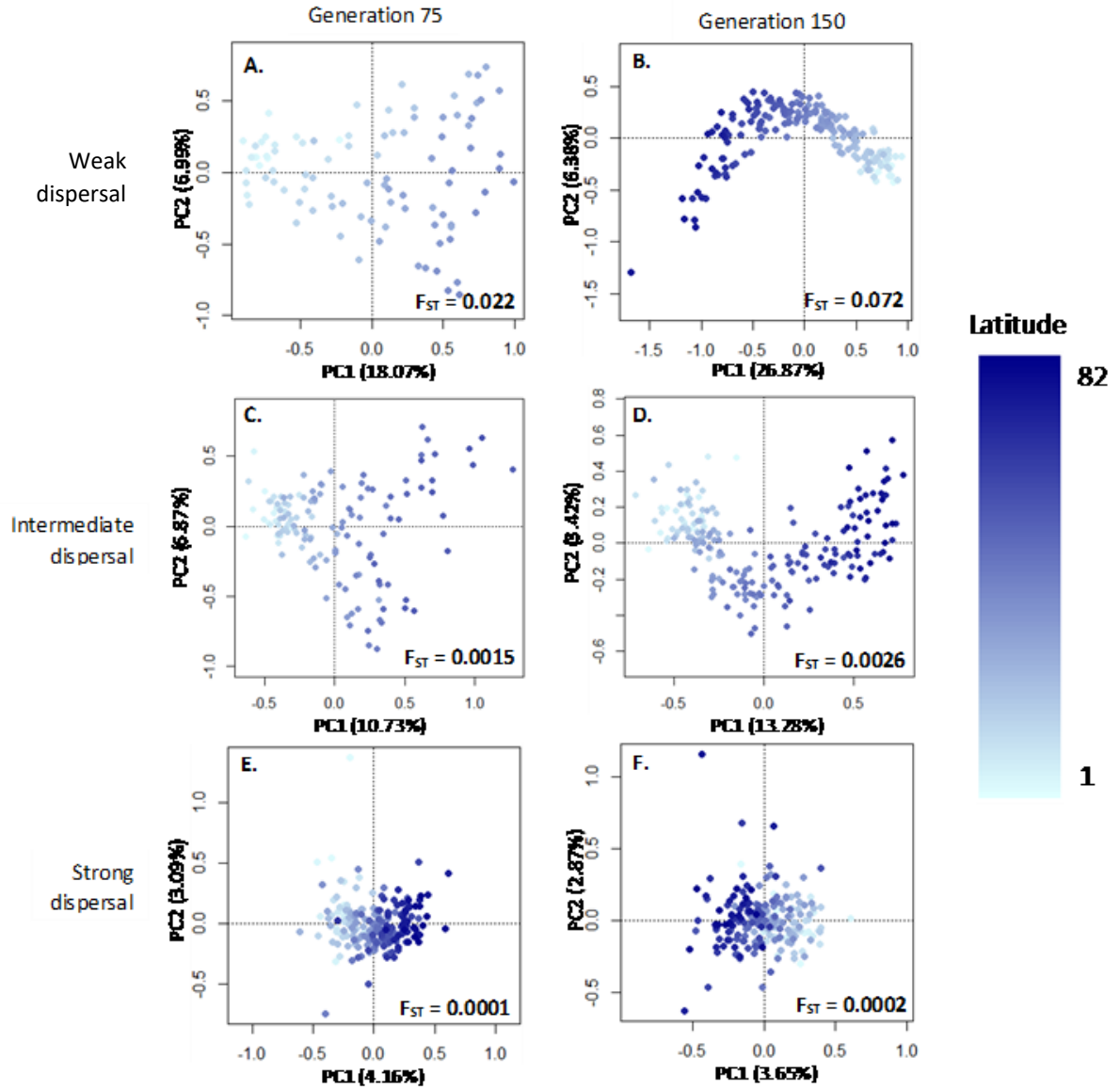


Figure 2: PCA plots of the genomes of individuals from 200 cells sampled across the simulated grid showing the genetic structure of a population at a specific time step. These six plots (A to F) are based on three distinct simulations varying in dispersal capacity and at two different time steps. Blue gradient represents the latitudinal position of each cell: darker points are located in the upper end of the grid and vice versa. At the bottom right corner of each plot is the global F_{ST} value of the sampled population. Number in parentheses on axis indicates the proportion of variance explained by the principal component.

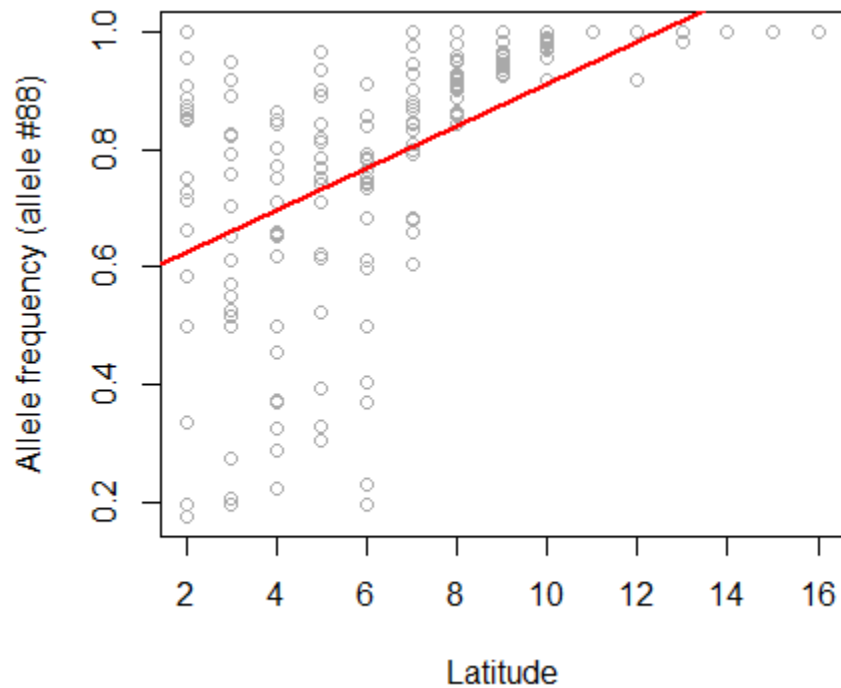


Figure 3: Example of a scatter plot illustrating the relationship between allele frequency and geographic position. Here data is shown for the first allele of neutral locus #88 from a simulation with weak dispersal capacity and intermediate strength of selection at generation 50 (early on during a range expansion). Points represent the relative allele frequencies in each cell (location X, Y) over the extent of the colonized area. The red line is the fitted line from the linear regression model. The slope of this line is then compared with the slope from the reference adaptive allele.

Identifying outlier loci

We observed a large number of neutral loci that showed slopes greater than the adaptive threshold for all levels of dispersal capacity and strength of selection (Figure 4 and Figure 5A-O). Neutral alleles were more likely to be identified as adaptive under conditions of weak dispersal, with early sampling relative to the start of the expansion, and when the adaptive locus used as a

reference for adaptation was under weak selection pressure (Figure 5A,D,G). We also observed in most scenarios a strong increase of those outlier loci following the end of the range expansion, owing to a phenomenon of adaptive allele fixation (more details below).

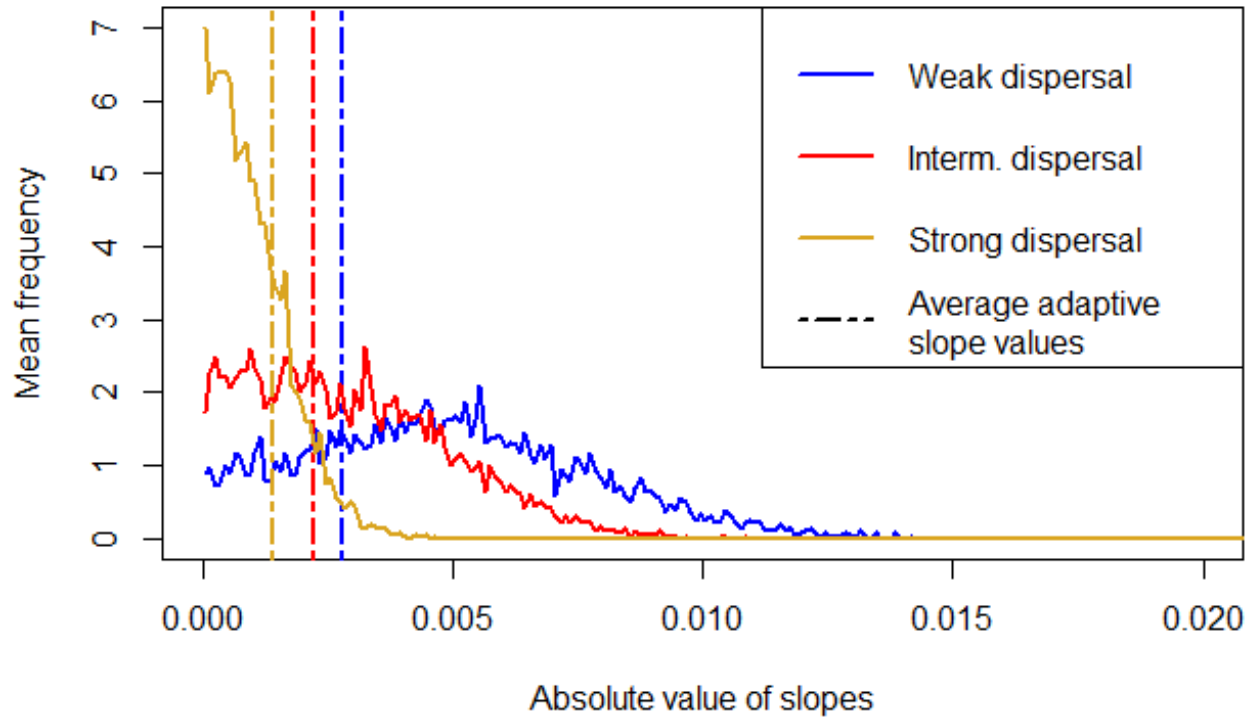
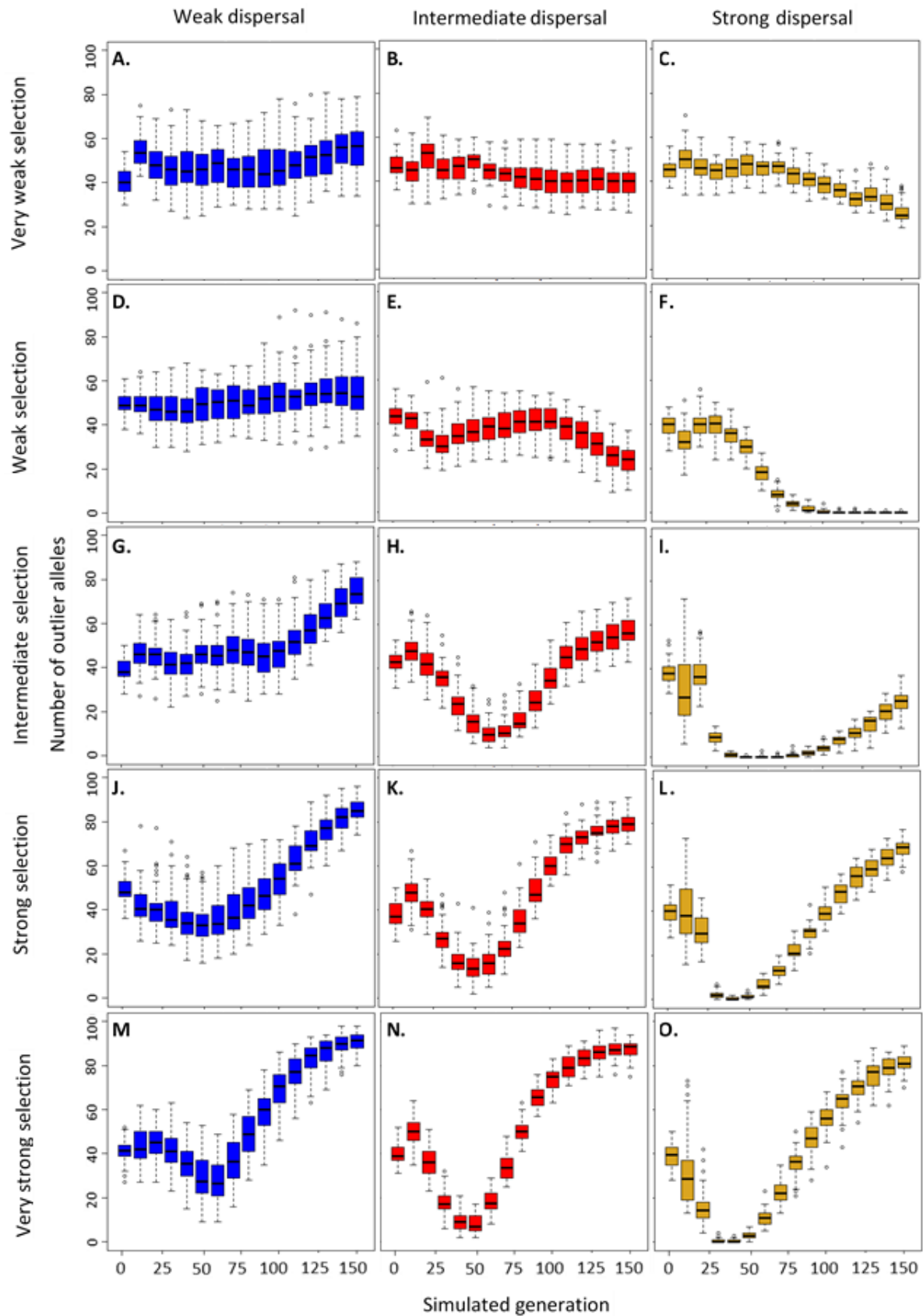


Figure 4: Frequency distribution of absolute neutral allele slopes for three different dispersal capacities at generation 150. Vertical dashed lines represent the respective mean allele frequency for the single simulated adaptive locus. Neutral loci with slope values greater than the adaptive threshold are more likely to be identified as potentially under selection. Depiction of all neutral allele slopes from all simulations with an intermediate strength of selection.

Figure 5 (Next page): Boxplots of neutral alleles with a higher slope value than the mean adaptive slope value, through time and for the three dispersal capacities and five strengths of selection. Letters A to O correspond to the data from each of the 15 scenarios. Colors distinguish the three dispersal capacities.



Most scenarios exhibit similar changes through time with respect to the proportion of loci whose slopes exceed that of the simulated adaptive locus (i.e., “outliers”) (Figure 5A-O). At the beginning of the range expansions (i.e., generations 0 to ~30), a high proportion (~45 %) of neutral loci exceed the adaptive threshold. Around the 50th generation, this proportion decreases to a minimum of 20 to 40 % for a weak dispersal capacity (Figure 5J,M), 10 % for an intermediate dispersal capacity (Figure 5H,K,N) and around 0 % for a strong dispersal capacity (Figure 5F,I,L,O). Finally, it increases once again (Figure 5H-O), due to the adaptive locus becoming fixed in the population (see below). Weaker strengths of selection do not show such a pattern, the number of “outlier” loci decreasing with higher dispersal capacities without increasing back (Figure 5C,E,F). Weak dispersal results in a higher proportion of “outlier” neutral loci, no matter the timing of sampling (Figure 5A,D,G,J,M). Furthermore, we note that the decrease in the proportion of “outlier” loci occurs earlier and more suddenly in scenarios with higher dispersal (Figure 5F,H,I,K,L,N,O).

All three factors examined had a significant influence on the risk of misclassifying neutral loci as adaptive (high F-ratios and p-value < 0,001; Table II). We also identified a significant three-way interaction among dispersal, selection, and timing (Table II). Although all factors and interactions were significant, it is informative to examine the F statistics of each (Nakagawa & Cuthill, 2007). F-ratios thus give a better impression of which variables or interactions are more important, because larger values indicate a greater inconsistency with the null hypothesis of no relationship between variables. Timing of sampling came out as the most influential variable, followed by dispersal, and strength of selection. The interaction between strength of selection and timing of sampling was also quite influential. For the other interactions, which mostly have smaller F-ratios, strength of selection x dispersal capacity was the most important, followed by the interaction between the three variables and finally the interaction between dispersal and timing of sampling.

Parameters	F value	P value
Selection (S)	531,344	<0,001
Dispersal (D)	2331,173	<0,001
Timing of sampling (T)	3548,137	<0,001
S x D	139,549	<0,001
S x T	1557,207	<0,001
D x T	90,319	<0,001
S x D x T	92,217	<0,001

Table II. Summary of the ANOVA results, which indicates significative effects of strengths of selection, dispersal capacity, timing of sampling and their interactions on the proportion of neutral loci being misidentified as adaptive.

Fixation of the adaptive locus

The adaptive locus became fixed in every simulation, generally near the end of the simulated expansion. This fixation affected the proportion of neutral loci classified as potentially adaptive because fixation creates slopes that are close to zero due to the absence of spatial variation in allele frequencies (Figure 6). The rate at which fixation occurs varies among scenarios and accounts for the upward trajectories of all curves in Figure 5. Specifically, fixation of the adaptive locus occurs more quickly under strong than weak selection (Figure 5H,K,N). Once the adaptive locus is fixed and the reference slope is equal to zero, almost all neutral alleles are identified as adaptive using our simple detection method (Figure 2).

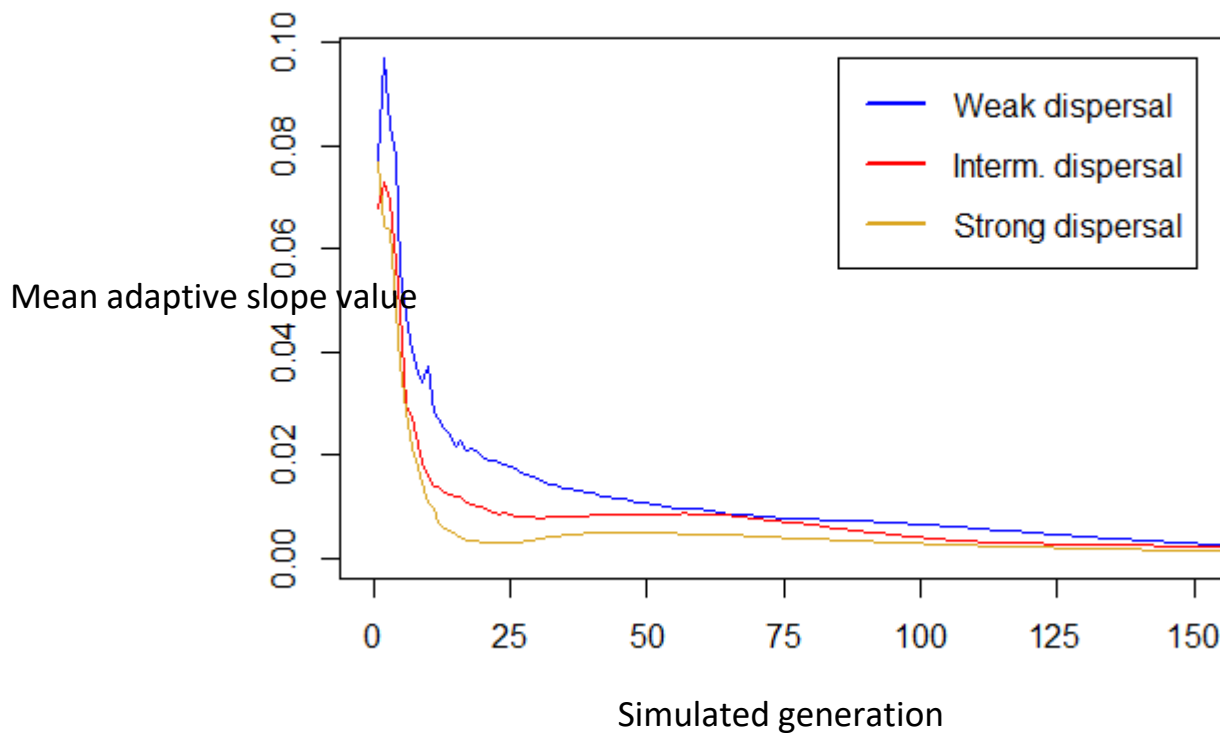


Figure 6: Mean adaptive slope over 50 replicates through time for three different dispersal capacities and at an intermediate strength of selection. Data from other scenarios exhibit similar patterns.

Discussion

We investigated the dynamics of range expansion and its effects on our ability to identify loci under selection using the mountain pine beetle as a case study. We employed a novel, spatially explicit individual-based simulation modelling approach (CDmetaPOP: Landguth et al., 2016) to simulate a series of range expansions driven by demographic and genetic processes that depicts accurately some of the main processes and genetic outcomes present in natural systems. We found that dispersal, strength of selection, and timing of sampling during the course of range expansion, all had a significant influence on the risk of misclassifying neutral loci as adaptive.

Dispersal

Spatial patterns from neutral loci are less likely to be confounded with spatial patterns from adaptive loci in organisms with strong dispersal capacity. Greater dispersal results in greater genetic connectivity that prevents local differentiation and neutral clines (Bialozyt et al., 2006; Polechová & Barton, 2015), and reduces the likelihood of allele surfing (Excoffier & Ray, 2008). In contrast, reduced dispersal and limited gene flow facilitates allele surfing (Excoffier et al., 2009) and creates neutral clines that can be confounded with adaptive patterns.

Differentiating spatial patterns of neutral loci from those of adaptive loci becomes more complex when considering a population which uses a combination of dispersal distances, as it is the case of the MPB (L. Safranyik et al., 1992) and our simulations. In expanding systems, an essential step to assess the dispersal related risk of mistaking a neutral locus as adaptive would be to verify the rate of expansion to determine if a mode of dispersal is dominant (Kot, Lewis, & Van den Driessche, 1996). However, it has previously been observed that long-distance dispersal has a greater impact than short-distance dispersal on the development of spatial genetic structure (Neubert & Caswell, 2000), as it could be observed in our strong dispersal simulations, where there was only a small proportion of long-distance dispersal events. Characterizing the MPB range expansion would allow the distinction of areas colonized by short- or long-distance dispersal events (or both), supporting more robust interpretation of neutral and adaptive genetic structures.

Timing of sampling

Early sampling during a range expansion is more likely to produce a higher rate of neutral loci exhibiting spatial patterns of adaptive loci. With fewer individuals, populations are more sensitive to genetic drift early in the expansion process (Mallet, 2001). Moreover, as the expansion progresses, the initial genetic signal of range expansion, including stronger spatial genetic structure and a greater proportion of “outlier” loci, decays (Lotterhos & Whitlock, 2015). This can be explained by the expansion front reaching the opposite end of the landscape, allowing a greater homogenization of genetic material through gene flow. Therefore, when trying to identify adaptive loci, it is important to take into consideration the current progress of the expansion. Spatial variation in genetic diversity and expected heterozygosity can be used to determine the age of a range expansion (Swaegers et al., 2015). According to our results, when

sampling a short-distance dispersing population undergoing a range expansion since at least 40 generations, neutral loci are less susceptible to mimic the spatial patterns of adaptive loci. On the other hand, if long-distance dispersal is present, sampling from 25 generations since the beginning of the expansion would allow low false discovery rates. In all cases, we found that timing of sampling influences the probability of incorrectly identifying a neutral locus as adaptive.

Selection

We found that weaker strengths of selection were responsible for lower local adaptation and thus weaker adaptive genetic signature (i.e., small slopes values) across the landscapes (Yeaman & Whitlock, 2011), resulting in a higher false discovery rate. This rate can also be affected by the assumptions regarding the spatial model of selection assumed and the strength of that selection (Forester et al., 2016). The strength of selection affected the rate at which the simulated adaptive allele became fixed and hence affected the temporal dynamics of our identification of loci potentially under selection. The effect of fixation on the calculation of the false discovery rate was most pronounced in scenarios with higher strengths of selection and high dispersal capacities. In those scenarios, it took fewer generations for the fixed allele to spread across the landscape than with the weak selective surface and weak dispersal capacity. Furthermore, simulations with weaker strengths of selection did not experienced that strong fixation phenomenon. In fact, with the weak selection scenarios (very weak and weak strengths of selection), the total length of the simulations was probably too short to allow the spread of the fixed adaptive allele across the landscape.

Models of selection

The model of selection we used assumed that selective coefficient increases gradually along a linear latitudinal gradient, in the style of a temperature gradient (Keller, Alexander, Holderegger, & Edwards, 2013). In this selection model, individuals are assumed to leave well-suited habitat to new areas with less suitable environmental conditions. This model was chosen because it best captures the hypothesized evolutionary dynamics of range expansion in MPB outbreaks. For example, we know that MPB populations continue to expand north- and east-ward, where mean temperatures are generally lower, which reduces larvae winter survival rates (Cullingham et al., 2011; L. Safranyik & Linton, 1991).

Although conceptually appropriate, use of this selection model resulted in the challenge of adaptive locus fixation. Fixation results in a uniform, high allele frequency across the entire simulation study area and a slope very close to zero (Figure 6). Because classification of a neutral locus as adaptive depends on whether its slope exceeds that of the reference adaptive locus, fixation creates the unrealistic situation in which all neutral loci appear “adaptive” as they all have a slope greater than zero. An additional challenge presented by this emergent property is that the rate at which the adaptive locus reaches fixation varies as a function of the strength of selection and the timing of sampling during the range expansion (Figure 5G-O). Indeed, with stronger strengths of selection, it took less time to reach fixation across the landscape. Moreover, a single adaptive locus was simulated in every replicate of each scenario for use as a reference. Our simple method for classifying neutral loci as potentially adaptive was based on comparison with this single adaptive locus. Including several adaptive loci in such models would better represent the complexity of real genetic data (Hoban, 2014). Future work in this area would benefit from further critical examination of the different spatial models of selection and the ways in which they affect spatial patterns in adaptive alleles. Such models could include antagonistic pleiotropy, which promotes balancing selection and prevent an allele from becoming extinct or fixed (Jones et al., 2013). The model could also include non-continuous selection surfaces, typical of heterogeneous landscapes (Gauffre et al., 2015; Ortego, Riordan, Gugger, & Sork, 2012). Even when considering a temperature gradient, a more realistic two dimensions model could have also included heterogeneity along the west-east axis of the grid. However, the simple selective linear gradient used in this paper could be transposed to the current MPB outbreak, as temperature gradient is one of the main mortality agents in expanding populations (Bentz et al., 2011).

Conclusions

This study improves our understanding of the specific conditions which neutral loci may mimic the spatial patterns of adaptive loci during a range expansion. We found that a combination of weak dispersal capacity, early timing of sampling, and weak selective pressure promotes a higher false discovery rate. Thus, it is advised to investigate *a priori* the conditions under which range expansion is occurring to better recognize the potential pitfalls of naive analyses. Although improved knowledge of species dispersal, selective pressures and their spatial configuration, expansion history will provide a better impression of the probability of mis-

identifying neutral loci as adaptive, we are not yet at a point where we can explicitly include this information in landscape genomic methods. Clearly, many challenges remain to addressing the problem of neutral genetic variation mimicking adaptive one as a result of range expansion. Landscape heterogeneity and linkage disequilibrium are a few examples of other processes affecting the performance of landscape genomics tools. Until these questions are resolved, and new methods become available, great care must be taken in interpreting the results of these tools when populations are undergoing range expansion.

Acknowledgements

We acknowledge funding for this research under the TRIA-Net project (<http://www.thetriaproject.ca>) from the Natural Science and Engineering Council of Canada's (NSERC) Strategic Network Grants program. This project was also supported by scholarships from the Forest Complexity Modelling (FCM) NSERC CREATE program to PM. We also acknowledge the assistance of Calcul Québec and Compute Canada for use of their supercomputing infrastructure (supercomputer *Colosse* from Université Laval à Québec). Finally, we thank Erin Landguth for valuable exchanges regarding CDmetaPOP.

Supporting information

Table S.1: Simulation parameters summary. Parameters not mentioned are either set to default value or unused.

A. *PatchVars* input file: with a total of 1804 rows, one for each cell.

Parameters	Value	Justification
Patch	1 to 1804	Id of each simulated cell
X	1 to 22	x-coordinate
Y	1 to 82	y-coordinate
K	30	Cell carrying capacity
N0	30 for 4 cells ; 0 for the rest of them	Initial population of 120 individuals in the South portion of the grid
Natal.Grounds	1	Individuals can occupy a natal cell
Genes.Initialize	random	Random initial assignment of alleles
Mortality.Back	100 for border cells ; N for the rest of them	Mortality associates with individuals at the border cells: they can't disperse out of those cells (they die)
Migration	0	Individuals can't disperse during simulator's emigration phase
Straying	1	Individuals disperse during simulator's immigration phase
Fitness_AA	100 for border cells ; 0 for the rest	Gradients of mortality following our spatial selection variation model
Fitness_Aa	100 for border cells ; then 0 to 15 (weak), 0 to 25 (intermediate) or 0 to 35	

	(strong selection) for the rest	
Fitness_aa	100 for border cells ; then 0 to 30 (weak), 0 to 50 (intermediate) or 0 to 70 (strong selection) for the rest	

B. *ClassVars* input file: with 3 rows, one for each age class.

Parameters	Value	Justification
Age class	0, 1 and 2	Three age classes in order to produce a univoltin species model in CDmetaPOP
Distribution	1 for Age 0	Simulations begin with only Age 0 individuals
Percent Female	66 for all age classes	Sex ratio for MPB is about 1 male for 2 females, although it can change according to climatic conditions and epidemic stages (Les Safranyik & Wilson, 2006)
Age Mortality Out %	100 for Age 3	All Age 2 individuals are discarded each generation during the simulation process (they don't contribute to dispersal or reproduction)
Age Mortality Back %	100 for Age 3	
Migration	0 for all age classes	Age 1 individuals are dispersing during the immigration phase, before reproducing
Straying	1 for Age 1	

Male Maturation	1 for Age 1	Age 1 individuals become sexually mature right before the simulator's reproduction process
Female Maturation	1 for Age 1	
Fecundity_mu	60 for Age 1	MPB females lay on average 60 eggs per mating event (Les Safranyik & Wilson, 2006)

C. PopVars input file

Parameters	Value	Justification
runtime	900*	Simulations spanning to three times the required duration to fill the grid (with a weak dispersal capacity)
output_years	2	Extract genetic data at the beginning of every generation
cdclimgentime	0 2	One burn-in generation (i.e. 2 time steps) at the beginning of a simulation, where the cost distance matrix only allows dispersal within the 4 cells initially occupied
cdmat (mate, dispout, dispback and stray)	Patchclosed.csv Cdmatrix.csv	
Dispersal parameters	<i>See Figure 1</i>	
Freplace	N	Males MPB sometimes reproduce with more than one female (Les Safranyik & Wilson, 2006)
Mreplace	Y	
offno	2	The number of offspring in a clutch follows a Poisson distribution law
eggFrequency	1	Fertilized MPB females always lay eggs

Egg_Mortality	88	A mean value of mortality taken from the different life stages of MPB, from larvae to adulthood (Amman & Cole, 1983; L. Safranyik, Barclay, Thomson, & Riel, 1999; L. Safranyik et al., 2010; Les Safranyik & Wilson, 2006)
Egg_FemalePercent	66	Sex ratio for MPB is about 1 male for 2 females, although it can change according to climatic conditions and epidemic stages (Les Safranyik & Wilson, 2006)
loci	101	100 neutral loci and 1 adaptive locus
alleles	2	2 alleles for each locus (bi-allelic)
muterate	0,0005	Default mutation model
mutationtype	random	
cdevolveans	1	Selection model (1 adaptive locus)

*Although we ran all simulations for that duration, we only kept data up to time step 300 for my analyses. This corresponds roughly to the time when range expansions in weak dispersal scenarios is completed and when the process of adaptive fixation becomes more prevalent thorough the landscape.

Chapitre 3 : Conclusion

Dans le cadre de ce projet, je me suis intéressé à l'effet de la capacité de dispersion, de la force de sélection et du temps durant une expansion de territoire sur notre capacité d'identifier des loci adaptatifs avec une approche de génomique du paysage. J'ai construit un modèle de simulation d'expansion de territoire qui consiste en une simplification de la situation épidémique actuelle du dendroctone du pin dans l'ouest canadien et qui inclut des phénomènes génétiques complexes. J'ai fait varier la capacité de dispersion des individus selon trois niveaux (distance maximale faible, intermédiaire et forte) ainsi que la force de sélection du paysage selon cinq niveaux (très faible, faible, intermédiaire, forte et très forte) et j'ai testé leur effet sur notre capacité à distinguer les loci neutres de ceux adaptatifs.

Mes résultats ont démontré que la faible capacité de dispersion entraîne un plus haut taux de faux-positifs, en raison de la plus forte structure génétique qui est maintenue à travers le territoire. Inversement, lorsque les événements de dispersion à longue distance sont présents, moins d'allèles neutres démontrant un patron spatial similaire à celui des allèles adaptatifs. J'explique ce résultat par une plus grande homogénéisation génétique de la population (c.-à-d. à la fois en périphérie qu'au cœur de l'aire de répartition), ce qui réduit les différenciations génétiques engendrées par la dérive et autres phénomènes génétiques. En effet, une capacité de dispersion limitée entraîne un plus grand impact du surf d'allèles, qui produit des gradients de fréquences alléliques neutres.

J'ai démontré que chez les espèces en expansion de territoire, comme les espèces envahissantes, il est important de prendre en considération la complexité du mode de dispersion lorsque des méthodes d'identification de loci adaptatifs sont utilisées. Dans le contexte du dendroctone du pin, dont la capacité de dispersion est complexe et inclut à la fois des dispersions à courtes et à très longues distances, mes résultats indiquent que l'utilisation d'une approche de génomique du paysage dépend de l'échelle particulière et du contexte de l'étude. En effet, si l'échantillonnage génétique ne comprend que des populations de dendroctones fondées suite à la dispersion à courte distance, il faudra alors faire attention au phénomène de surf d'allèles et interpréter avec soin les résultats des méthodes de génomique du paysage.

Les différentes forces de sélection ont eu un impact plus faible sur notre capacité de distinguer les allèles neutres de ceux adaptatifs. En effet, une faible sélection n'était associée qu'à une légère augmentation du taux de faux-positifs. Toutefois, mon modèle de sélection a eu des répercussions majeures sur ma méthode d'identification des allèles neutres potentiellement considérés comme adaptatifs. En effet, un processus de fixation de l'allèle adaptatif non délétère prédominant vers la fin de chacune de mes expansions de territoire a fait en sorte que le taux de faux-positifs fut artificiellement gonflé. Quoique conceptuellement adéquat, le modèle de sélection ne permet pas ainsi de tirer des conclusions robustes sur une partie de mes simulations. Explorer d'autres modèles de sélection, comme un modèle de pléiotropie antagoniste (c.-à-d., un modèle où un allèle est avantageé sur une partie du territoire et désavantageé sur l'autre), permettrait d'éviter ce problème de fixation (Jones et al., 2013). Actuellement, plusieurs informations sur la survie du dendroctone du pin nous permettent d'évaluer sa susceptibilité quant à certaines conditions de son environnement tels les basses températures, l'humidité sous l'écorce, la présence de prédateurs/parasitoïdes et la capacité de défense des hôtes (Les Safranyik & Wilson, 2006). Avec ces renseignements, il est possible d'estimer les forces de sélection à l'œuvre dans un habitat donné, dans l'optique d'avoir une idée générale des surfaces de sélection qui affectent les populations de dendroctones. Au final, ceci permettra d'adapter les interprétations des résultats des méthodes de génomique du paysage : si la surface de sélection est faible, il faut être plus prudent.

Enfin, j'ai également testé comment le temps depuis le début de l'expansion, j'en conclue qu'un échantillonnage précoce de la population en expansion est susceptible de conduire à une plus grande proportion de faux-positifs. Les conditions initiales de l'expansion sont marquées par une forte implication de la dérive génétique, qui restera importante tout au long de l'expansion du territoire. Néanmoins, avec le temps et la progression de l'expansion, ce taux de faux-positifs va décroître, permettant une interprétation plus robuste des résultats des méthodes de génomiques du paysage. Ceci s'explique par le rôle homogénéisant du flux génique, dont l'importance augmente avec le temps. En effet, plus les gènes se répartissent à travers le territoire, plus les signaux génétiques d'expansion du territoire présents dans les loci neutres (c.-à-d. des gradients de fréquences alléliques) tenderont à disparaître. Ainsi, la dispersion doit être prise en considération, car mes résultats montrent que lorsqu'une dispersion à longue distance est présente, il suffit d'un peu plus de 20 générations pour que le

taux de faux-positifs soit assez bas. Lorsque ce mode de dispersion est absent, c'est plus de 40 générations qui sont nécessaires avant que l'homogénéisation du matériel génétique permette de réduire le taux de faux-positifs. À la lumière de ces résultats, considérant le jeune âge de l'expansion actuelle du dendroctone du pin (environ une quinzaine de générations; Les Safranyik & Wilson, 2006) et sa capacité de dispersion à longues distances, il faut s'attendre à observer des patrons génétiques spatiaux marqués pour les loci neutres, ce qui gonfle le taux de faux-positifs des méthodes de détection de loci adaptatifs.

À la lumière de ces résultats, j'ai plusieurs conseils et directives à fournir aux études utilisant des méthodes de génomique du paysage pour identifier des loci potentiellement adaptatifs chez des populations en expansion de territoire. Connaître à priori le contexte démographique de l'expansion de territoire permettra d'interpréter avec plus de robustesse les résultats de telles méthodes. Ainsi, il est important d'évaluer le mode de dispersion de l'espèce à l'étude dominant la progression de l'expansion et d'évaluer l'importance d'événements de dispersion à longue distance, si présente. Également, l'âge de l'expansion est un autre paramètre important à examiner : depuis combien de générations la population est-elle en expansion? Enfin, malgré les difficultés que ça comporte, estimer les forces de sélection qui agissent sur le fitness des individus à travers le territoire colonisé permettra de faire des interprétations moins erronées des résultats des méthodes de génomique du paysage.

Plusieurs améliorations peuvent encore être apportées au modèle utilisé, notamment un raffinement du modèle de sélection (c.-à-d. autres surfaces de sélection plus réalistes), l'ajout de paramètres génétiques importants comme le déséquilibre de liaison (« *linkage disequilibrium* »), l'inclusion de plus d'un locus adaptatif par simulation et l'hétérogénéité du paysage. Ce dernier élément nécessiterait l'élaboration de surfaces de mortalité qui correspondent aux paysages naturels rencontrés par les populations en expansion, tenant compte de la fragmentation des habitats et des structures du paysage qui facilitent la dispersion. Tous ces facteurs sont susceptibles de complexifier l'identification adéquate de loci adaptatifs à l'aide de méthodes de génomique du paysage en renforçant la structure génétique particulière de la population (François et al., 2016). Ainsi, l'inclusion de ces éléments dans le modèle serait bénéfique pour accroître nos connaissances sur cette problématique. Enfin, appliquer directement les méthodes de génomique du paysage sur nos données permettrait de directement comparer leurs performances et identifier avec plus d'exactitude les conditions

sous lesquelles il devient difficile de différencier la variation génétique neutre de celle adaptative.

Avec les changements climatiques et les activités humaines qui perturbent les écosystèmes de la planète, les années à venir connaîtront des altérations importantes sur la distribution de nombreuses espèces (Chuang & Peterson, 2016). Les travaux futurs dans le domaine seront particulièrement importants pour raffiner les outils qui nous permettront d'étudier ces espèces envahissantes ou irruptives ou encore celles qui suivent simplement leur enveloppe climatique changeante (Hansen et al., 2012). Les méthodes actuellement utilisées en génomique du paysage nécessitent des améliorations essentielles pour tenir compte de la structure génétique générée sur des paysages complexes chez une population qui n'est pas à l'équilibre démographique (Lotterhos & Whitlock, 2015). Identifier les loci adaptatifs avec plus de certitude est donc important pour les chercheurs qui tentent de découvrir de nouveaux gènes impliqués dans le succès d'une espèce envahissante ou qui souhaitent valider le caractère adaptatif des gènes précédemment étudiés (François et al., 2016). Par conséquent, les résultats de ce projet de recherche contribuent à mener une interprétation plus prudente des résultats de méthodes de génomique du paysage.

Bibliographie

- Adams, A. S., Aylward, F. O., Adams, S. M., Erbilgin, N., Aukema, B. H., Currie, C. R., . . . Raffa, K. F. (2013). Mountain Pine Beetles Colonizing Historical and Naïve Host Trees Are Associated with a Bacterial Community Highly Enriched in Genes Contributing to Terpene Metabolism. *Applied and Environmental Microbiology*, 79(11), 3468-3475.
- Addison, A. L., Powell, J. A., Six, D. L., Moore, M., & Bentz, B. J. (2013). The role of temperature variability in stabilizing the mountain pine beetle–fungus mutualism. *Journal of Theoretical Biology*, 335, 40-50. doi: <http://dx.doi.org/10.1016/j.jtbi.2013.06.012>
- Amman, G. D., & Cole, W. E. (1983). Mountain Pine Beetle Dynamics in Lodgepole Pine Forests - Part II: Population Dynamics. *Forest Service*, 63.
- Austerlitz, F., Jung-Muller, B., Godelle, B., & Gouyon, P.-H. (1997). Evolution of coalescence times, genetic diversity and structure during colonization. *Theoretical Population Biology*, 51, 148-164.
- Barrett, S. C. H. (2015). Foundations of invasion genetics: the Baker and Stebbins legacy. *Molecular Ecology*, 24, 1927-1941. doi: 10.1111/mec.13014
- Beaumont, M. A., & Balding, D. J. (2004). Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, 13, 969-980. doi: 10.1111/j.1365-294X.2004.02125.x
- Bell, G., & Gonzalez, A. (2011). Adaptation and evolutionary rescue in metapopulations experiencing environmental deterioration. *Science*, 332(6035), 1327-1330.
- Bentz, B. J., Bracewell, R. R., Mock, K. E., & Pfrender, M. E. (2011). Genetic architecture and phenotypic plasticity of thermally-regulated traits in an eruptive species, *Dendroctonus ponderosae*. *Evolutionary Ecology*, 25, 1269-1288. doi: 10.1007/s10682-011-9474-x
- Bialozyt, R., Ziegenhagen, B., & Petit, R. J. (2006). Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. *Journal of Evolutionary Biology*, 19, 12-20.
- Bock, D. G., Caseys, C., Cousens, R. D., Hahn, M. A., Heredia, S. M., Hübner, S., . . . Rieseberg, L. H. (2015). What we still don't know about invasion genetics. *Molecular Ecology*, 24, 2277-2297. doi: 10.1111/mec.13032
- Bohonak, A. J. (1999). Dispersal, Gene Flow, and Population Structure. *The Quarterly Review of Biology*, 74(1), 21-45.
- Boone, C., White, J. C., Wulder, M. A., Robertson, C., & Nelson, T. A. (2013). Impact of Forest Fragmentation on Patterns of Mountain Pine Beetle-Caused Tree Mortality. *Forests*, 4, 279-295. doi: 10.3390/f4020279
- Bracewell, R. R., Pfrender, M. E., Mock, K. E., & Bentz, B. J. (2013). Contrasting Geographic Patterns of Genetic Differentiation in Body Size and Development Time with Reproductive Isolation in *Dendroctonus ponderosae* (Coleoptera: Curculionidae, Scolytinae). *Annals of the Entomological Society of America*, 106(3), 385-391.
- Braschler, B., & Hill, J. K. (2007). Role of larval host plants in the climate-driven range expansion of the butterfly *Polygonia c-album*. *Journal of Animal Ecology*, 76, 415-423.
- Bridle, J. R., Polechová, J., Kawata, M., & Butlin, R. K. (2010). Why is adaptation prevented at ecological margins? New insights from individual-based simulations. *Ecology Letters*, 13, 485-494. doi: 10.1111/j.1461-0248.2010.01442.x
- Brown, G. P., Phillips, B. L., & Shine, R. (2015). Directional dispersal has not evolved during the cane toad invasion. *Functional Ecology*, 29, 830-838. doi: 10.1111/1365-2435.12397

- Burton, O. J., & Travis, J. M. J. (2008a). The frequency of fitness peak shifts is increased at expanding range margins due to mutation surfing. *Genetics*, 179, 941-950. doi: 10.1534/genetics.108.087890
- Burton, O. J., & Travis, J. M. J. (2008c). Landscape structure and boundary effects determine the fate of mutations occurring during range expansions. *Heredity*, 101, 329-340.
- Cahill, A. E., & Levinton, J. S. (2016). Genetic differentiation and reduced genetic diversity at the northern range edge of two species with different dispersal modes. *Molecular Ecology*, 25, 515-526. doi: 10.1111/mec.13497
- Castillo, J. A., Epps, C. W., Davis, A. R., & Cushman, S. A. (2014). Landscape effects on gene flow for a climate-sensitive montane species, the American pika. *Molecular Ecology*, 23, 843-856. doi: 10.1111/mec.12650
- Chen, G.-B., Lee, S. H., Zhu, Z.-X., Benyamin, B., & Robinson, M. R. (2016). EigenGWAS: finding loci under selection through genome-wide association studies of eigenvectors in structures populations. *Heredity*, 117, 51-61.
- Chen, H., Ott, P., Wang, J., & Ebata, T. (2014). A positive response of mountain pine beetle to pine forest- clearcut edges at the landscape scale in British Columbia, Canada. *Landscape Ecology*, 29, 1625-1639. doi: 10.1007/s10980-014-0090-3
- Chen, H., & Walton, A. (2011). Mountain pine beetle dispersal: spatiotemporal patterns and role in the spread and expansion of the present outbreak. *Ecosphere*, 2(6), art66. doi: 10.1890/ES10-00172.1
- Chuang, A., & Peterson, C. R. (2016). Expanding population edges: theories, traits, and trade-offs. *Global Change Biology*, 22, 494-512. doi: 10.1111/gcb.13107
- Clobert, J., Le Galliard, J.-F., Cote, J., Meylan, S., & Massot, M. (2009). Informed dispersal, heterogeneity in animal dispersal syndromes and the dynamics of spatially structured populations. *Ecology Letters*, 12, 197-209. doi: 10.1111/j.1461-0248.2008.01267.x
- Cooke, J. (2013). TRIA Network: Turning Risk Into Action for the Mountain Pine Beetle Epidemic. 83.
- Coop, G., Witonsky, D., Rienzo, A. D., & Pritchard, J. K. (2010). Using environmental correlations to identify loci underlying local adaptation. *Genetics*, 185, 1411-1423. doi: 10.1534/genetics.110.114819
- Coulon, A., Cosson, J.-F., Angibault, J. M., Cargnelutti, B., Galan, M., Morellet, N., . . . Hewison, A. J. M. (2004). Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual-based approach. *Molecular Ecology*, 13, 2841-2850. doi: 10.1111/j.1365-294X.2004.02253.x
- Coulon, A., Guillot, G., Cosson, J.-F., Angibault, J. M. A., Aulagnier, S., Cargnelutti, B., . . . Hewison, A. J. M. (2006). Genetic structure is influenced by landscape features: empirical evidence from a roe deer population. *Molecular Ecology*, 15(1669-1679). doi: 10.1111/j.1365-294X.2006.02861.x
- Cristescu, M. E. (2015). Genetic reconstructions of invasion history. *Molecular Ecology*, 24, 2212-2225. doi: 10.1111/mec.13117
- Cudmore, T. J., Björklund, N., Carroll, A. L., & Lindgren, B. S. (2010). Climate change and range expansion of an aggressive bark beetle: evidence of higher beetle reproduction in naïve host tree populations. *Journal of Applied Ecology*, 47, 1036-1043. doi: 10.1111/j.1365-2664.2010.01848.x
- Cullingham, C. I., Cooke, J. E. K., Dang, S., Davis, C. S., Cooke, B. J., & Coltman, D. W. (2011). Mountain pine beetle host-range expansion threatens the boreal forest. *Molecular Ecology*, 20, 2157-2171.

- Cunningham, C. I., James, P. M. A., Cooke, J. E. K., & Coltman, D. W. (2012). Characterizing the physical and genetic structure of the lodgepole pine x jack pine hybrid zone: mosaic structure and differential introgression. *Evolutionary Applications*, 5, 879-891.
- Cushman, S. A., Shirk, A., & Landguth, E. L. (2012). Separating the effects of habitat area, fragmentation and matrix resistance on genetic differentiation in complex landscapes. *Landscape Ecology*, 27, 369-380. doi: 10.1007/s10980-011-9693-0
- Cushman, S. A., Shirk, A. J., & Landguth, E. L. (2013). Landscape genetics and limiting factors. *Conservation genetics*, 14, 263-274. doi: 10.1007/s10592-012-0396-0
- Cwynar, L. C., & MacDonald, G. M. (1987). Geographical variation of lodgepole pine in relation to population history. *The American Naturalist*, 129(3), 463-469.
- de la Giroday, H.-M. C., Carroll, A. L., & Aukema, B. H. (2012). Breach of the northern Rocky Mountain geoclimatic barrier: initiation of range expansion by the mountain pine beetle. *Journal of Biogeography*, 39, 1112-1123.
- de la Giroday, H.-M. C., Carroll, A. L., Lindgren, B. S., & Aukema, B. H. (2011). Incoming! Association of landscape features with dispersing mountain pine beetle populations during a range expansion event in western Canada. *Landscape Ecology*, 26, 1097-1110. doi: 10.1007/s10980-011-9628-9
- De Mita, S., Thuillet, A.-C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J., & Vigouroux, Y. (2013). Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*, 22, 1383-1399. doi: 10.1111/mec.12182
- De Villemereuil, P., Frichot, E., Bazin, E., François, O., & Gaggiotti, O. (2014). Genome scan methods against more complex models: when and how much should we trust them? *Molecular Ecology*, 23, 2006-2019. doi: 10.1111/mec.12705
- Débarre, F., Ronce, O., & Gandon, S. (2013). Quantifying the effects of migration and mutation on adaptation and demography in spatially heterogeneous environments. *Journal of Evolutionary Biology*, 26, 1185-1202. doi: 10.1111/jeb.12132
- Dileo, M. F., Rouse, J. D., Dávila, J. A., & Loughheed, S. C. (2013). The influence of landscape on gene flow in the eastern massasauga rattlesnake (*Sistrurus c. catenatus*): insight from computer simulations. *Molecular Ecology*, 22, 4483-4498. doi: 10.1111/mec.12411
- Dlugosch, K. M., & Parker, I. M. (2008). Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, 17, 431-449. doi: 10.1111/j.1365-294X.2007.03538.x
- Drury, K. L. S., Drake, J. M., Lodge, D. M., & Dwyer, G. (2007). Immigration events dispersed in space and time: factors affecting invasion success. *Ecological Modelling*, 206, 63-78. doi: 10.1016/j.ecolmodel.2007.03.017
- Duforet-Frebourg, N., Luu, K., Laval, G., Bazin, E., & Blum, M. G. B. (2015). Detecting genomic signatures of natural selection with principal component analysis: application to the 1000 Genomes data. arXiv:1504.04543v04545 [q-bio.PE].
- Endler, J. A. (1973). Gene flow and population differentiation. *Science*, 179(4070), 243-250.
- Epperson, B. K., McRae, B. H., Scribner, K., Cushman, S. A., Rosenberg, M. S., Fortin, M.-J., . . . Dale, M. R. T. (2010). Utility of computer simulations in landscape genetics. *Molecular Ecology*, 19, 3549-3564. doi: 10.1111/j.1365-294X.2010.04678.x
- Epps, C. W., & Keyghobadi, N. (2015). Landscape genetics in a changing world: disentangling historical and contemporary influences and inferring change. *Molecular Ecology*, 24, 6021-6040. doi: 10.1111/mec.13454

- Evenden, M. L., Whitehouse, C. M., & Sykes, J. (2014). Factors Influencing Flight Capacity of the Mountain Pine Beetle (Coleoptera: Curculionidae: Scolytinae). *Environmental Entomology*, 43(1), 187-196.
- Excoffier, L., Foll, M., & Petit, R. J. (2009). Genetic Consequences of Range Expansions. *Annual Review of Ecology, Evolution, and Systematics*, 40, 481-501. doi: 10.1146/annurev.ecolsys.39.110707.173414
- Excoffier, L., & Ray, N. (2008). Surfing during population expansions promotes genetic revolutions and structuration. *TRENDS in Ecology and Evolution*, 23(7), 347-351.
- Fabian, D. K., Kapun, M., Nolte, V., Kofler, R., Schmidt, P. S., Schlötterer, C., & Flatt, T. (2012). Genome-wide patterns of latitudinal differentiation among populations of *Drosophila melanogaster* from North America. *Molecular Ecology*, 21, 4748-4769. doi: 10.1111/j.1365-294X.2012.05731.x
- Fitzpatrick, S. W., Gerberich, J. C., Kronenberger, J. A., Angeloni, L. M., & Funk, W. C. (2015). Locally adapted traits maintained in the face of high gene flow. *Ecology Letters*, 18, 37-47.
- Flatt, T. (2016). Genomics of clinal variation in *Drosophila*: disentangling the interactions of selection and demography. *Molecular Ecology*, 25, 1023-1026.
- Ford, E. B. (1975). Ecological Genetics, 4th edition. *Chapman and Hall, London*.
- Forester, B. R., Jones, M. R., Joost, S., Landguth, E. L., & Lasky, J. R. (2016). Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Molecular Ecology*, 25, 104-120. doi: 10.1111/mec.13476
- François, O., Martins, H., Caye, K., & Schoville, S. D. (2016). Controlling false discoveries in genome scans for selection. *Molecular Ecology*, 25, 454-469. doi: 10.1111/mec.13513
- Frichot, E., Schoville, S., de Villemereuil, P., Gaggiotti, O., & François, O. (2015). Detecting adaptive evolution based on association with ecological gradients: Orientation matters! *Heredity*, 115, 22-28.
- Frichot, E., Schoville, S. D., Bouchard, G., & François, O. (2013). Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*.
- Garnier, J., Roques, L., & Hamel, F. (2012). Success rate of a biological invasion in terms of the spatial distribution of the founding population. *Bulletin of Mathematical Biology*, 74, 453-473. doi: 10.1007/s11538-011-9694-9
- Gauffre, B., Mallez, S., Chapuis, M.-P., Leblois, R., Litrico, I., Delaunay, S., & Badenhauer, I. (2015). Spatial heterogeneity in landscape structure influences dispersal and genetic structure: empirical evidence from a grasshopper in an agricultural landscape. *Molecular Ecology*, 24, 1713-1728. doi: 10.1111/mec.13152
- Gautier, M. (2015). Genome-wide scan for adaptive divergence and association with population-specific covariates. *Genetics*. doi: <http://dx.doi.org/10.1101/023721>
- Gehara, M., Summers, K., & Brown, J. L. (2013). Population expansion, isolation and selection: novel insights on the evolution of color diversity in the strawberry poison frog. *Evolutionary Ecology*, 27, 797-824. doi: 10.1007/s10682-013-9652-0
- Goodsman, D. W., Cooke, B., Coltman, D. W., & Lewis, M. A. (2014). The genetic signature of rapid range expansions: How dispersal, growth and invasion speed impact heterozygosity and allele surfing. *Theoretical Population Biology*, 98, 1-10.
- Goudet, J., & Jombart, T. (2015). hierfstat: Estimation and Tests of Hierarchical F-Statistics (Version R package version 0.04-22). Retrieved from <http://CRAN.R-project.org/package=hierfstat>

- Graciá, E., Botella, F., Anadón, J. D., Edelaar, P., Harris, D. J., & Giménez, A. (2013). Surfing in tortoises? Empirical signs of genetic structuring owing to range expansion. *Biology Letters*, 9, 20121091.
- Günther, T., & Coop, G. (2012). Robust identification of local adaptation from allele frequencies. *Genetics*, 195, 205-220.
- Haag, C. R., Saastamoinen, M., Marden, J. H., & Hanski, I. (2005). A candidate locus for variation in dispersal rate in a butterfly metapopulation. *Proceedings of the Royal Society B*, 272, 2449-2456. doi: 10.1098/rspb.2005.3235
- Haas, R. J., & Payseur, B. A. (2016). Fifteen years of genomewide scans for selection: trends, lessons and unaddressed genetic sources of complication. *Molecular Ecology*, 25, 5-23. doi: 10.1111/mec.13339
- Haddad, N. M., Brudvig, L. A., Clobert, J., Davies, K. F., Gonzalez, A., Holt, R. D., . . . Townshend, J. R. (2015). Habitat fragmentation and its lasting impact on Earth's ecosystems. *Science Advances*, 1(2), e1500052. doi: 10.1126/sciadv.1500052
- Hague, M. T. J., & Routman, E. J. (2016). Does population size affect genetic diversity? A test with sympatric lizard species. *Heredity*, 116, 92-98.
- Hallatschek, O., Hersen, P., Ramanathan, S., & Nelson, D. R. (2007). Genetic drift at expanding frontiers promotes gene segregation. *PNAS*, 104(50), 19926-19930.
- Hallatschek, O., & Nelson, D. (2008). Gene surfing in expanding populations. *Theoretical Population Biology*, 73(1), 158-170. doi: 10.1016/j.tpb.2007.08.008
- Hallatschek, O., & Nelson, D. R. (2009). Life at the front of an expanding population. *Evolution*, 64(1), 193-206. doi: 10.1111/j.1558-5646.2009.00809.x
- Hansen, M. M., Olivieri, I., Waller, D. M., & Nielsen, E. E. (2012). Monitoring adaptive genetic responses to environmental change. *Molecular Ecology*, 21, 1311-1329. doi: 10.1111/j.1365-294X.2011.05463.x
- Hartl, D. L., & Clark, A. G. (1997). Principles of population genetics (Vol. 116). *Sunderland: Sinauer associates*.
- Hartman, G. (1997). Notes on age at dispersal of beaver (*Castor fiber*) in an expanding population. *Revue canadienne de zoologie*, 75(6), 959-962.
- Hassall, C., Thompson, D. J., & Harvey, I. F. (2009). Variation in morphology between core and marginal populations of three British damselflies. *Aquatic Insects*, 31(3), 187-197. doi: 10.1080/01650420902776708
- Hill, J. K., Griffiths, H. M., & Thomas, C. D. (2011). Climate Change and Evolutionary Adaptations at Species' Range Margins. *Annual Review of Entomology*, 56, 143-159. doi: 10.1146/annurev-ento-120709-144746
- Hill, J. K., Thomas, C. D., & Blakeley, D. S. (1999). Evolution of flight morphology in a butterfly that has recently expanded its geographic range. *Oecologia*, 121(2), 165-170.
- Hoban, S. (2014). An overview of the utility of population simulation software in molecular ecology. *Molecular Ecology*, 23, 2383-2401. doi: 10.1111/mec.12741
- Hoban, S., Bertorelle, G., & Gaggiotti, O. E. (2012). Computer simulations: tools for population and evolutionary genetics. *Nature Reviews - Genetics*, 13, 110-122.
- Holderegger, R., Herrmann, D., Poncet, B. N., Gugerli, F., Thuiller, W., Taberlet, P., . . . Manel, S. (2008). Land ahead: Using genome scans to identify molecular markers of adaptive relevance. *Plant Ecology & Diversity*, 1(2), 273-283. doi: 10.1080/17550870802338420
- Holderegger, R., Kamm, U., & Gugerli, F. (2006). Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology*, 21, 797-807. doi: 10.1007/s10980-005-5245-9

- James, P. M. A., Coltman, D. W., Murray, B. W., Hamelin, R. C., & Sperling, F. A. H. (2011). Spatial Genetic Structure of a Symbiotic Beetle-Fungal System: Toward Multi-Taxa Integrated Landscape Genetics. *PLoS ONE*, 6(10), e25359. doi: 10.1371/journal.pone.0025359
- Janes, J. K., Li, Y., Keeling, C. I., Yuen, M. M. S., Boone, C. K., Cooke, J. E. K., . . . Sperling, F. A. H. (2014). How the Mountain Pine Beetle (*Dendroctonus ponderosae*) Breached the Canadian Rocky Mountains. *Molecular Biology and Evolution*, 31(7), 1803-1815. doi: 10.1093/molbev/msu135
- Jensen, J. D., Foll, M., & Bernatchez, L. (2016). The past, present and future of genomic scans for selection. *Molecular Ecology*, 25, 1-4.
- Jombart, T., Pontier, D., & Dufour, A.-B. (2009). Genetic markers in the playground of multivariate analysis. *Heredity*, 102, 330-341.
- Jones, M. R., Forester, B. R., Teufel, A. I., Adams, R. V., Anstett, D. N., Goodrich, B. A., . . . Manel, S. (2013). Integrating landscape genomics and spatially explicit approaches to detect loci under selection in clinal populations. *Evolution*, 67(12), 3455-3468. doi: 10.1111/evo.12237
- Joost, S., Bonin, A., Bruford, W., Després, L., Conord, C., Erhardt, G., & Taberlet, P. (2007). A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology*, 16, 3955-3969. doi: 10.1111/j.1365-294X.2007.03442.x
- Joost, S., Kalbermatten, M., & Bonin, A. (2008). Spatial analysis method (SAM): A software tool combining molecular and environmental data to identify candidate loci for selection. *Molecular Ecology Resources*, 8, 957-960. doi: 10.1111/j.1755-0998.2008.02162.x
- Joost, S., Vuilleumier, S., Jensen, J. D., Schoville, S., Leempoel, K., Stucki, S., . . . Manel, S. (2013). Uncovering the genetic basis of adaptive change: on the intersection of landscape genomics and theoretical population genetics. *Molecular Ecology*, 22, 3659-3665.
- Jordal, B. H. (2014). 3.7.12 Scolytinae Latreille, 1806. In B. Leschen (Ed.), *Arthropoda: Insecta: Coleoptera*. (Vol. Volume 3: Morphology and Systematics (Phytophaga), pp. 349-358). Berlin: De Gruyter.
- Kautz, M., Schopf, R., & Ohser, J. (2013). The "sun-effect": microclimatic alterations predispose forest edges to bark beetle infestations. *European Journal of Forest Research*, 132, 453-465. doi: 10.1007/s10342-013-0685-2
- Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, 7, 1225-1241. doi: 10.1111/j.1461-0248.2004.00684.x
- Keller, I., Alexander, J. M., Holderegger, R., & Edwards, P. J. (2013). Widespread phenotypic and genetic divergence along altitudinal gradients in animals. *Journal of Evolutionary Biology*, 26, 2527-2543. doi: 10.1111/jeb.12255
- Khatchikian, C. E., Prusinski, M. A., Stone, M., Bryon Backenson, P., Wang, I.-N., Foley, E., . . . Brisson, D. (2015). Recent and rapid population growth and range expansion of the Lyme disease tick vector, *Ixodes scapularis*, in North America. *Evolution*, 69(7), 1678-1689.
- Kilkenny, F. F., & Galloway, L. F. (2013). Adaptive divergence at the margin of an invaded range. *Evolution*, 67(3), 722-731. doi: 10.1111/j.1558-5646.2012.01829.x
- Kimura, M., & Crow, J. F. (1963). The measurement of effective population number. *Evolution*, 17(3), 279-288.
- Kirkpatrick, M., & Barton, N. H. (1997). Evolution of a Species' Range. *The American Naturalist*, 150(1), 1-23.

- Klopfstein, S., Currat, M., & Excoffier, L. (2006). The fate of mutations surfing on the wave of a range expansion. *Molecular Biology and Evolution*, 23(3), 482-490. doi: 10.1093/molbev/msj057
- Kot, M., Lewis, M. A., & Van den Driessche, P. (1996). Dispersal data and the spread of invading organisms. *Ecology*, 77(7), 2027-2042.
- Kurz, W. A., Dymond, C. C., Stinson, G., Rampley, G. J., Neilson, E. T., Carroll, A. L., . . . Safranyik, L. (2008). Mountain pine beetle and forest carbon feedback to climate change. *Nature - Letters*, 452, 987-990. doi: 10.1038/nature06777
- Kurz, W. A., Stinson, G., Rampley, G. J., Dymond, C. C., & Neilson, E. T. (2008). Risk of natural disturbances makes future contribution of Canada's forests to the global carbon cycle highly uncertain. *Proceedings of the National Academy of Sciences (USA)*, 105(5), 1551-1555.
- Landguth, E. L., & Balkenhol, N. (2012). Relative sensitivity of neutral versus adaptive genetic data for assessing population differentiation. *Conservation genetics*, 13, 1421-1426. doi: 10.1007/s10592-012-0354-x
- Landguth, E. L., Bearlin, A., Day, C., & Dunham, J. (2016). CDMetaPOP: an eco-evolutionary metapopulation simulation model for population viability analysis in landscape and riverscape genetics. v1.0.
- Landguth, E. L., & Cushman, S. A. (2010). CDPOP: A spatially explicit cost distance population genetics program. *Molecular Ecology Resources*, 10, 156-161.
- Landguth, E. L., Cushman, S. A., & Johnson, N. A. (2012). Simulating natural selection in landscape genetics. *Molecular Ecology Resources*, 12, 363-368. doi: 10.1111/j.1755-0998.2011.03075.x
- Landguth, E. L., Fedy, B. C., Oyler-McCance, S. J., Garey, A. L., Emel, S. L., Mumma, M., . . . Cushman, S. A. (2012). Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern. *Molecular Ecology Resources*, 12, 276-284.
- Landguth, E. L., Muhlfeld, C. C., & Luikart, G. (2012). CDFISH: an individual-based, spatially-explicit, landscape genetics simulator for aquatic species in complex riverscape. *Conservation Genetics Resources*, 4, 133-136. doi: 10.1007/s12686-011-9492-6
- Larroque, J., Ruetz, S., Vandel, J.-M., & Devillard, S. (2016). Divergent landscape effects on genetic differentiation in two populations of the European pine marten (*Martes martes*). *Landscape Ecology*, 31, 517-531. doi: 10.1007/s10980-015-0281-6
- Lehe, R., Hallatschek, O., & Peliti, L. (2012). The rate of beneficial mutations surfing on the wave of a range expansion. *PLoS Computational Biology*, 8(3), e1002447. doi: 10.1371/journal.pcbi.1002447
- Liebl, A. L., & Martin, L. B. (2012). Exploratory behaviour and stressor hyper-responsiveness facilitate range expansion of an introduced songbird. *Proceedings of the Royal Society B*, 279, 4375-4381. doi: 10.1098/rspb.2012.1606
- Lieutier, F., Yart, A., & Salle, A. (2009). Stimulation of tree defenses by Ophiostomatoid fungi can explain attack success of bark beetles on conifers. *Annals of Forest Science*, 66, 801. doi: 10.1051/forest/2009066
- Ling, S. D., Johnson, C. R., Frusher, S., & King, C. K. (2008). Reproductive potential of a marine ecosystem engineer at the edge of a newly expanded range. *Global Change Biology*, 14, 907-915. doi: 10.1111/j.1365-2486.2008.01543.x
- Liu, H., Prugnolle, F., Manica, A., & Balloux, F. (2006). A geographically explicit genetic model of worldwide human-settlement history. *The American Journal of Human Genetics*, 79, 230-237.

- Logan, J. A., & Bentz, B. J. (1999). Model analysis of mountain pine beetle (Coleoptera: Scolytidae) seasonality. *Environmental Entomology*, 28(6), 924-934.
- Logan, J. A., Régnière, J., & Powell, J. A. (2003). Assessing the impacts of global warming on forest pest dynamics. *Frontiers in Ecology and the Environment*, 1(3), 130-137.
- Lombaert, E., Estoup, A., Facon, B., Joubard, B., Grégoire, J.-C., Jannin, A., . . . Guillemaud, T. (2014). Rapid increase in dispersal during range expansion in the invasive ladybird *Harmonia axyridis*. *Journal of Evolutionary Biology*, 27, 508-517. doi: 10.1111/jeb.12316
- Lotterhos, K. E., & Whitlock, M. C. (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, 24, 1031-1046. doi: 10.1111/mec.13100
- Mallet, J. (2001). Gene Flow. *Insect Movement: Mechanisms and Consequences*, 337-360.
- Manel, S., & Holderegger, R. (2013). Ten years of landscape genetics. *TRENDS in Ecology and Evolution*, 28(10), 614-621. doi: <http://dx.doi.org/10.1016/j.tree.2013.05.012>
- Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. *TRENDS in Ecology and Evolution*, 18(4), 189-197.
- Maruyama, T., & Fuerst, P. A. (1984). Population bottlenecks and nonequilibrium models in population genetics. I. Allele numbers when populations evolve from zero variability. *Genetics*, 108, 745-763.
- Meirmans, P. G. (2012). The trouble with isolation by distance. *Molecular Ecology*, 21, 2839-2846.
- Moran, E. V., & Alexander, J. M. (2014). Evolutionary responses to global change: lessons from invasive species. *Ecology Letters*, 17, 637-649. doi: 10.1111/ele.12262
- Moreau, C., Bhérer, C., Vézina, H., Jomphe, M., Labuda, D., & Excoffier, L. (2011). Deep human genealogies reveal a selective advantage to be on an expanding wave front. *Science*, 334, 1148-1150.
- Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews*, 82, 591-605. doi: 10.1111/j.1469-185X.2007.00027.x
- Nathan, R., Perry, G., Cronin, J. T., Strand, A. E., & Cain, M. L. (2003). Methods for estimating long-distance dispersal. *OIKOS*, 103, 261-273.
- Nei, M. (1973a). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences USA*, 70(12), 3321-3323.
- Nei, M. (1973c). Analysis of Gene Diversity in Subdivided Populations. *Proceedings of the National Academy of Sciences*, 70(12), 3321-3323.
- Neubert, M. G., & Caswell, H. (2000). Demography and dispersal: Calculation and sensitivity analysis of invasion speed for structured populations. *Ecology*, 81(6), 1613-1628.
- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., . . . Wagner, H. H. (2016). *vegan: Community Ecology Package* (Version R package version 2.3-5). Retrieved from <http://CRAN.R-project.org/package=vegan>
- Ortego, J., Riordan, E. C., Gugger, P. F., & Sork, V. L. (2012). Influence of environmental heterogeneity on genetic diversity and structure in an endemic southern Californian oak. *Molecular Ecology*, 21(3210-3223). doi: 10.1111/j.1365-294X.2012.05591.x
- Oyler-McCance, S. J., DeYoung, R. W., Fike, J. A., Hagen, C. A., Johnson, J. A., Larsson, L. C., & Patten, M. A. (2016). Rangewide genetic analysis of Lesser Prairie-Chicken reveals population structure, range expansion, and possible introgression. *Conservation genetics*, 17, 643-660. doi: 10.1007/s10592-016-0812-y

- Pariset, L., Joost, S., Ajmone Marsan, P., & Valentini, A. (2009). Landscape genomics and biased FST approaches reveal single nucleotide polymorphisms under selection in goat breeds of North-East Mediterranean. *BMC Genetics*, 10(7), 1-8. doi: 10.1186/1471-2156-10-7
- Parisod, C., & Holderegger, R. (2012). Adaptive landscape genetics: pitfalls and benefits. *Molecular Ecology*, 21, 3644-3646.
- Pedersen, L. (2003). How serious is the mountain pine beetle problem? From a timber supply perspective. In *Proceedings of the Mountain Pine Beetle Symposium: Challenges and Solutions*, 10-20.
- Pejchar, L., & Mooney, H. A. (2009). Invasive species, ecosystem services and human well-being. *TRENDS in Ecology and Evolution*, 24(9), 497-504.
- Peter, B. M., & Slatkin, M. (2013). Detecting range expansions from genetic data. *Evolution*, 67(11), 3274-3289.
- Peter, B. M., & Slatkin, M. (2015). The effective founder effect in a spatially expanding population. *Evolution*, 69(3), 721-734.
- Phillips, B. L., Brown, G. P., & Shine, R. (2010). Life-history evolution in range-shifting populations. *Ecology*, 91(6), 1617-1627.
- Pitman, G. B., & Vité, J. P. (1969). Aggregation behavior of *Dendroctonus ponderosae* (Coleoptera: Scolytidae) in response to chemical messengers. *The Canadian Entomologist*, 101(2), 143-149. doi: <http://dx.doi.org/10.4039/Ent101143-2>
- Polechová, J., & Barton, N. H. (2015). Limits to adaptation along environmental gradients. *PNAS*, 112(20), 6401-6406.
- Preisler, H. K., Hicke, J. A., Ager, A. A., & Hayes, J. L. (2012). Climate and weather influences on spatial temporal patterns of mountain pine beetle populations in Washington and Oregon. *Ecology*, 93(11), 2421-2434.
- Raby, G. D., Gutowsky, L. F. G., & Fox, M. G. (2010). Diet composition and consumption rate in round goby (*Neogobius melanostomus*) in its expansion phase in the Trent River, Ontario. *Environmental Biology of Fishes*, 89, 143-150. doi: 10.1007/s10641-010-9705-y
- Raffa, K. F., Aukema, B. H., Bentz, B. J., Carroll, A. L., Hicke, J. A., Turner, M. G., & Romme, W. H. (2008). Cross-scale Drivers of Natural Disturbances Prone to Anthropogenic Amplification: The Dynamics of Bark Beetle Eruptions. *BioScience*, 58(6), 501-517.
- Raffa, K. F., Powell, E. N., & Townsend, P. A. (2013). Temperature-driven range expansion of an irruptive insect heightened by weakly coevolved plant defenses. *PNAS*, 110(6), 2193-2198. doi: 10.1073/pnas.1216666110
- Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, 24, 4348-4370. doi: 10.1111/mec.13322
- Reznick, D., Bryant, M. J., & Bashey, F. (2002). r- and K-selection revisited: the role of population regulation in life-history evolution. *Ecology*, 83(6), 1509-1520.
- Richardson, J. L., Brady, S. P., Wang, I. J., & Spear, S. F. (2016). Navigating the pitfalls and promise of landscape genetics. *Molecular Ecology*, 25, 849-863.
- Robertson, C., Nelson, T. A., Jelinski, D. E., Wulder, M. A., & Boots, B. (2009). Spatial-temporal analysis of species range expansion: the case of the mountain pine beetle, *Dendroctonus ponderosae*. *Journal of Biogeography*, 36, 1446-1458. doi: 10.1111/j.1365-2699.2009.02100.x
- Roe, A. D., James, P. M. A., Rice, A. V., Cooke, J. E. K., & Sperling, F. A. H. (2011). Spatial community structure of mountain pine beetle fungal symbionts across a latitudinal gradient. *Microbial Ecology*, 62(2), 347-360. doi: 10.1007/s00248-011-9841-8

- Rohlf, F. J., & Schnell, G. D. (1971). An investigation of the isolation-by-distance model. *The American Naturalist*, 105(944), 295-324.
- Rolshausen, G., Muttalib, S., Kaeuffer, R., Oke, K. B., Hanson, D., & Hendry, A. P. (2015). When maladaptive gene flow does not increase selection. *Evolution*, 69(9), 2289-2302. doi: 10.1111/evo.12739
- Roulin, A. C., Mariadassou, M., Hall, M. D., Walser, J.-C., Haag, C., & Ebert, D. (2015). High genetic variation in resting-stage production in a metapopulation: is there evidence for local adaptation? *Evolution*, 69(10), 2747-2756. doi: 10.1111/evo.12770
- Rózsa, J., Strand, T. M., Montadert, M., Kozma, R., & Höglund, J. (2016). Effects of a range expansion on adaptive and neutral genetic diversity in dispersal limited Hazel grouse (*Bonasa bonasia*) in the French Alps. *Conservation genetics*, 17, 401-412. doi: 10.1007/s10592-015-0792-3
- Safranyik, L., Barclay, H., Thomson, A., & Riel, W. G. (1999). A population dynamics model for the mountain pine beetle, *Dendroctonus ponderosae* Hopk. (Coleoptera : Scolytidae). 386.
- Safranyik, L., Carroll, A. L., Régnière, J., Langor, D. W., Riel, W. G., Shore, T. L., . . . Taylor, S. W. (2010). Potential for Range Expansion of Mountain Pine Beetle into the Boreal Forest of North America. *The Canadian Entomologist*, 142(5), 415-442.
- Safranyik, L., & Linton, D. A. (1991). Unseasonably low fall and winter temperatures affecting mountain pine beetle and pine engraver beetle populations and damage in the British Columbia Chilcotin Region. *Journal of the Entomological Society of British Columbia*, 88, 17-21.
- Safranyik, L., Linton, D. A., Silversides, R., & McMullen, L. H. (1992). Dispersal of released mountain pine beetles under the canopy of a mature lodgepole pine stand. *Journal of Applied Entomology*, 113, 441-450.
- Safranyik, L., & Wilson, B. (2006). The Mountain Pine Beetle: A Synthesis of Biology, Management, and Impacts on Lodgepole Pine. *Natural Resources Canada*, 304.
- Sakai, A. K., Allendorf, F. W., Holt, J. S., Lodge, D. M., Molofsky, J., With, K. A., . . . Weller, S. G. (2001). The population biology of invasive species. *Annual Review of Ecology, Evolution, and Systematics*, 32, 305-332.
- Samarasekera, G. D. N. G., Bartell, N. V., Lindgren, B. S., Cooke, J. E. K., Davis, C. S., James, P. M. A., . . . Murray, B. W. (2012). Spatial genetic structure of the mountain pine beetle (*Dendroctonus ponderosae*) outbreak in western Canada: historical patterns and contemporary dispersal. *Molecular Ecology*, 21, 2931-2948.
- Sanford, E., Holzman, S. B., Haney, R. A., Rand, D. M., & Bertness, M. D. (2006). Larval tolerance, gene flow, and the northern geographic range limit of fiddler crabs. *Ecology*, 87(11), 2882-2894.
- Schwartz, M. K., Luikart, G., McKelvey, K. S., & Cushman, S. A. (2010). Landscape genomics: a brief perspective. *Spatial Complexity, Informatics, and Wildlife Conservation*, 165-174. doi: 10.1007/978-4-431-87771-4_9
- Schweizer, R. M., Vonholdt, B. M., Harrigan, R., Knowles, J. C., Musiani, M., Coltman, D., . . . Wayne, R. K. (2016). Genetic subdivision and candidate genes under selection in North American grey wolves. *Molecular Ecology*, 25, 380-402. doi: 10.1111/mec.13364
- Sequeira, A. S., Normark, B. B., & Farrell, B. D. (2000). Evolutionary assembly of the conifer fauna: distinguishing ancient from recent associations in bark beetles. *Proceedings of the Royal Society of London B*, 267, 2359-2366. doi: 10.1098/rspb.2000.1292
- Siol, M., Wright, S. I., & Barrett, S. C. H. (2010). The population genomics of plant adaptation. *New Phytologist*, 188, 313-332. doi: 10.1111/j.1469-8137.2010.03401.x

- Slatkin, M., & Barton, N. H. (1989). A comparison of three indirect methods for estimating average levels of gene flow. *Evolution*, 43(7), 1349-1368.
- Slatkin, M., & Excoffier, L. (2012). Serial founder effects during range expansion: a spatial analog of genetic drift. *Genetics*, 191, 171-181. doi: 10.1534/genetics.112.139022
- Song, S. V., Downes, S., Parker, T., Oakeshott, J. G., & Robin, C. (2015). High nucleotide diversity and limited linkage disequilibrium in *Helicoverpa armigera* facilitates the detection of a selective sweep. *Heredity*, 115, 460-470. doi: 10.1038/hdy.2015.53
- Sousa, V., Peischl, S., & Excoffier, L. (2014). Impact of range expansions on current human genomic diversity. *Current Opinion in Genetics & Development*, 29, 22-30.
- Spear, S. F., Peterson, C. R., Matocq, M. D., & Storfer, A. (2005). Landscape genetics of the blotched tiger salamander (*Ambystoma tigrinum melanostictum*). *Molecular Ecology*, 14, 2553-2564. doi: 10.1111/j.1365-294X.2005.02573.x
- Storey, J. D., & Tibshirani, R. (2003). Statistical significance for genomewide studies. *PNAS*, 100(16), 9440-9445.
- Swaegers, J., Mergeay, J., van Geystelen, A., Therry, L., Larmuseau, M. H. D., & Stoks, R. (2015). Neutral and adaptive genomic signatures of rapid poleward range expansion. *Molecular Ecology*, 24, 6163-6176. doi: 10.1111/mec.13462
- Taberlet, P., Fumagalli, L., Wust-Saucy, A.-G., & Cosson, J.-F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, 7, 453-464.
- Therry, L., Lefevre, E., Bonte, D., & Stoks, R. (2014). Increased activity and growth rate in the non-dispersive aquatic stage of a damselfly at an expanding range edge. *Freshwater Biology*, 59, 1266-1277. doi: 10.1111/fwb.12346
- Thornton, P. K., Ericksen, P. J., Herrero, M., & Challinor, A. J. (2014). Climate variability and vulnerability to climate change: a review. *Global Change Biology*, 20, 3313-3328. doi: 10.1111/gcb.12581
- Tiffin, P., & Ross-Ibarra, J. (2014). Advances and limits of using population genetics to understand local adaptation. *TRENDS in Ecology and Evolution*, 29(12), 673-680. doi: <http://dx.doi.org/10.1016/j.tree.2014.10.004>
- Tigano, A., & Friesen, V. L. (2016). Genomics of local adaptation with gene flow. *Molecular Ecology*, 25, 2144-2164. doi: 10.1111/mec.13606
- Travis, J. M. J., Münkemüller, T., Burton, O. J., Best, A., Dytham, C., & Johst, K. (2007). Deleterious mutations can surf to high densities on the wave front of an expanding population. *Molecular Biology and Evolution*, 24(10), 2334-2343. doi: 10.1093/molbev/msm167
- van Strien, M. J., Holderegger, R., & Van Heck, H. J. (2014). Isolation-by-distance in landscapes: considerations for landscape genetics. *Heredity*, 1-11.
- Vasemägi, A. (2006). The Adaptive Hypothesis of Clinal Variation Revisited: Single-Locus Clines as a Result of Spatially Restricted Gene Flow. *Genetics*, 173, 2411-2414.
- Volis, S., Ormanbekova, D., & Shulgina, I. (2016). Role of selection and gene flow in population differentiation at the edge vs. interior of the species range differing in climatic conditions. *Molecular Ecology*, 25, 1449-1464. doi: 10.1111/mec.13565
- Wasserman, T. N., Cushman, S. A., Shirk, A. S., Landguth, E. L., & Littell, J. S. (2012). Simulating the effects of climate change on population connectivity of American marten (*Martes americana*) in the northern Rocky Mountains, USA. *Landscape Ecology*, 27, 211-225. doi: 10.1007/s10980-011-9653-8

- Watterson, G. A. (1986). The homozygosity test after a change in population size. *Genetics*, 112(899-907).
- Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 38(6), 1358-1370.
- Willi, Y., Van Buskirk, J., & Hoffman, A. A. (2006). Limits to the adaptive potential of small populations. *Annual Review of Ecology, Evolution, and Systematics*, 37, 433-458. doi: 10.1146/annurev.ecolsys.37.091305.110145
- Wilson, M. C., Chen, X.-Y., Corlett, R. T., Didham, R. K., Ding, P., Holt, R. D., . . . Yu, M. (2016). Habitat fragmentation and biodiversity conservation: key findings and future challenges. *Landscape Ecology*, 31, 219-227. doi: 10.1007/s10980-015-0312-3
- Wood, S. (1982). The bark and ambrosia beetles of North and Central America (Coleoptera: Scolytidae), a taxonomic monograph. *Salt Lake City: Brigham Young University*.
- Wood, S. L. (1963). A revision of the bark beetle genus *Dendroctonus* Erichson (Coleoptera: Scolytidae). *The Great Basin Naturalist*, 23(1-2), 1-117.
- Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics*, 15, 323-354.
- Yeaman, S., & Whitlock, M. C. (2011). The genetic architecture of adaptation under migration-selection balance. *Evolution*, 65(7), 1897-1911. doi: 10.1111/j.1558-5646.2011.01269.x
- Zenni, R. D., & Hoban, S. (2015). Loci under selection during multiple range expansions of an invasive plant are mostly population-specific, but patterns are associated with climate. *Molecular Ecology*, 24(13), 3360-3371.